155

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds
(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649A-5
Perfect score: 49
Sequence: 1 LSVRLGXPVKE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24626/app_query.fasta_1.1393
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649_@CGN_1_1_77_@runat_21012003_093151_24626
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	47	95.9	2682	9	US-09-891-332A-1	Sequence 1, Appli
	2	37	75.5	696	10	US-09-974-300-7373	Sequence 7373, Ap
	3	36	73.5	93	10	US-09-922-261-33	Sequence 33, Appl
c	4	36	73.5	464	10	US-09-967-768A-154	Sequence 154, App
	5	36	73.5	1096	10	US-09-922-261-26	Sequence 26, Appl
	6	36	73.5	43804	10	US-09-970-711-1	Sequence 1, Appli
	7	35	71.4	1074	9	US-10-004-717-18	Sequence 18, Appl
c	8	34	69.4	1945	10	US-09-731-872-37	Sequence 37, Appl
c	9	34	69.4	1953	10	US-09-731-872-41	Sequence 41, Appl
c	10	34	69.4	1969	10	US-09-731-872-59	Sequence 59, Appl
	11	34	69.4	2001	10	US-09-815-242-9539	Sequence 9539, Ap
	12	34	69.4	2031	10	US-09-815-242-9415	Sequence 9415, Ap
c	13	34	69.4	4668	10	US-09-286-240-5	Sequence 5, Appli
c	14	34	69.4	4668	10	US-09-863-777-1	Sequence 1, Appli
c	15	34	69.4	4668	10	US-09-880-107-2239	Sequence 2239, Ap
c	16	34	69.4	9388	9	US-09-924-400-141	Sequence 141, App
c	17	34	69.4	9388	10	US-09-810-936-141	Sequence 141, App
c	18	34	69.4	9388	10	US-09-429-755-141	Sequence 141, App
c	19	33	67.3	229	10	US-09-960-352-13394	Sequence 13394, A
	20	33	67.3	257	10	US-09-878-574-5560	Sequence 5560, Ap
c	21	33	67.3	257	10	US-09-878-574-6474	Sequence 6474, Ap
	22	33	67.3	273	10	US-09-878-574-8764	Sequence 8764, Ap
	23	33	67.3	277	10	US-09-867-701-6945	Sequence 6945, Ap
c	24	33	67.3	390	10	US-09-960-352-4518	Sequence 4518, Ap
	25	33	67.3	409	10	US-09-960-352-14105	Sequence 14106, A
	26	33	67.3	868	10	US-09-770-445-567	Sequence 567, App
c	27	33	67.3	891	9	US-09-938-842A-1994	Sequence 1994, Ap
c	28	33	67.3	2580	9	US-09-738-626-2964	Sequence 2964, Ap
	29	33	67.3	2768	10	US-09-905-983-4	Sequence 4, Appli
	30	33	67.3	2768	10	US-09-905-983-6	Sequence 6, Appli
c	31	33	67.3	8546	10	US-09-070-927A-146	Sequence 146, App
	32	33	67.3	19206	10	US-09-764-869-1258	Sequence 1258, Ap
	33	33	67.3	19274	10	US-09-764-869-1254	Sequence 1254, Ap
	34	33	67.3	31348	10	US-09-764-869-1259	Sequence 1259, Ap
	35	33	67.3	659158	9	US-09-771-208-20	Sequence 20, Appl
c	36	32	65.3	197	10	US-09-923-876-2960	Sequence 2960, Ap
c	37	32	65.3	222	10	US-09-960-352-4917	Sequence 4917, Ap
c	38	32	65.3	297	10	US-09-294-093B-322	Sequence 322, App
	39	32	65.3	363	10	US-09-770-791-477	Sequence 477, App
c	40	32	65.3	366	9	US-10-079-623-7	Sequence 7, Appli
c	41	32	65.3	428	10	US-09-960-352-4817	Sequence 4817, Ap
	42	32	65.3	490	10	US-09-917-800A-1063	Sequence 1063, Ap
	43	32	65.3	510	10	US-09-783-590-652	Sequence 652, App
	44	32	65.3	716	9	US-09-938-842A-4318	Sequence 4318, Ap
c	45	32	65.3	798	9	US-09-813-453A-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-891-332A-1
; Sequence 1, Application US/09891332A
; Patent No. US20020168646A1
; GENERAL INFORMATION:
; APPLICANT: Chatterjee, Deb K.
; Solus, Joseph
; Yang, Shuwei
; TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
; Nucleic Acid Fragments and Uses Thereof
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/891,332A
; FILING DATE: 27-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

```
;             APPLICATION NUMBER: 09/019,160
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 60/037,393
;             FILING DATE: 07-FEB-1997
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Esmond, Robert W.
;             REGISTRATION NUMBER: 32,893
;             REFERENCE/DOCKET NUMBER: 0942.4250002
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 202-371-2600
;             TELEFAX: 202-371-2540
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2682 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: both
;             TOPOLOGY: both
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-891-332A-1

Alignment Scores:
Pred. No.:              0.365          Length:        2682
Score:                  47.00          Matches:       10
Percent Similarity:     90.91%         Conservative:  0
Best Local Similarity:  90.91%         Mismatches:    1
Query Match:            95.92%         Indels:        0
DB:                     9              Gaps:          0

US-09-823-649A-5 (1-11) x US-09-891-332A-1 (1-2682)

Qy      1 LeuSerValArgLeuGly***ProValLysGlu 11
          ||||||||||||||||||||*  ||||||||||||
Db   2221 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 2253

RESULT 2
US-09-974-300-7373
; Sequence 7373, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7373
;   LENGTH: 696
;   TYPE: DNA
;   ORGANISM: Bacillus clausii
US-09-974-300-7373

Alignment Scores:
Pred. No.:              13.6           Length:        696
Score:                  37.00          Matches:       7
Percent Similarity:     90.00%         Conservative:  2
Best Local Similarity:  70.00%         Mismatches:    1
Query Match:            75.51%         Indels:        0
DB:                     10             Gaps:          0

US-09-823-649A-5 (1-11) x US-09-974-300-7373 (1-696)

Qy      2 SerValArgLeuGly***ProValLysGlu 11
          |||:::|||||||||   |||:::|||||
Db    657 TCCATGCGCCTCGGTTTGCCGTTAAAAGAA 686

RESULT 3
US-09-922-261-33
; Sequence 33, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
;  APPLICANT: COGENT NEUROSCIENCE, Inc.
;  APPLICANT:  Lo, Donald C.
;  APPLICANT:  Barney, Shawn
;  APPLICANT:  Thomas, Mary Beth
;  APPLICANT:  Portbury, Stuart D.
;  APPLICANT:  Puranam, Kasturi
;  APPLICANT:  Katz, Lawrence C.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
;  TITLE OF INVENTION:  CELL DEATH
;  FILE REFERENCE: 10001-005-999
;  CURRENT APPLICATION NUMBER: US/09/922,261
;  CURRENT FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: US/09/461,697
;  PRIOR FILING DATE: 1999-12-14
;  NUMBER OF SEQ ID NOS: 466
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
;   LENGTH: 93
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-922-261-33

Alignment Scores:
Pred. No.:              2.35           Length:        93
Score:                  36.00          Matches:       6
Percent Similarity:     90.00%         Conservative:  3
Best Local Similarity:  60.00%         Mismatches:    1
Query Match:            73.47%         Indels:        0
DB:                     10             Gaps:          0

US-09-823-649A-5 (1-11) x US-09-922-261-33 (1-93)

Qy      1 LeuSerValArgLeuGly***ProValLys 10
          :::|||:::|||||||||   |||||||:::
Db      1 ATGAGCATACGACTGGGGAGCCCACTGGAC 30

RESULT 4
US-09-967-768A-154/c
; Sequence 154, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
;  APPLICANT: Augustus, Meena
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sig
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-72
;  CURRENT APPLICATION NUMBER: US/09/967,768A
;  CURRENT FILING DATE:  2001-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,109
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,034
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US/60/236,111
;  PRIOR FILING DATE: 2000-09-28
;  NUMBER OF SEQ ID NOS: 325
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 154
;   LENGTH: 464
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(464)
;   OTHER INFORMATION: n=a,t,g or c
US-09-967-768A-154

Alignment Scores:
Pred. No.:              14.4           Length:        464
```

```
Score:                  36.00          Matches:        7
Percent Similarity:     81.82%         Conservative:   2
Best Local Similarity:  63.64%         Mismatches:     2
Query Match:            73.47%         Indels:         0
DB:                     10             Gaps:           0

US-09-823-649A-5 (1-11) x US-09-967-768A-154 (1-464)

Qy       1 LeuSerValArgLeuGly***ProValLysGlu 11
             |||   |||||||:::|||   ||||||||||:::
Db     311 CTTCCTGTTCGAATGGGTTATCCAGTAAAAAAG 279

RESULT 5
US-09-922-261-6
; Sequence 26, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
;  APPLICANT: COGENT NEUROSCIENCE, Inc.
;  APPLICANT:  Lo, Donald C.
;  APPLICANT:  Barney, Shawn
;  APPLICANT:  Thomas, Mary Beth
;  APPLICANT:  Portbury, Stuart D.
;  APPLICANT:  Puranam, Kasturi
;  APPLICANT:  Katz, Lawrence C.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
;  TITLE OF INVENTION:  CELL DEATH
;  FILE REFERENCE: 10001-005-999
;  CURRENT APPLICATION NUMBER: US/09/922,261
;  CURRENT FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: US/09/461,697
;  PRIOR FILING DATE: 1999-12-14
;  NUMBER OF SEQ ID NOS: 456
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
;   LENGTH: 1096
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-922-261-26

Alignment Scores:
Pred. No.:              37.9           Length:         1096
Score:                  36.00          Matches:        6
Percent Similarity:     90.00%         Conservative:   3
Best Local Similarity:  60.00%         Mismatches:     1
Query Match:            73.47%         Indels:         0
DB:                     10             Gaps:           0

US-09-823-649A-5 (1-11) x US-09-922-261-26 (1-1096)

Qy       1 LeuSerValArgLeuGly***ProValLys 10
             :::|||:::|||||||||   |||||||:::
Db     249 ATGAGCATACGACTGGGGAGCCCAGTGGAG 278

RESULT 6
US-09-970-711-1
; Sequence 1, Application US/09970711
; Patent No. US20020081279A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Adam
;  APPLICANT:  Cotten, Matthew
;  APPLICANT:  Chiocca, Susanna
;  APPLICANT:  Kurzbauer, Robert
;  APPLICANT:  Schaffner, Gotthold
;  TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
;  FILE REFERENCE: 0652.1800001
;  CURRENT APPLICATION NUMBER: US/09/970,711
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/171,461
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: PCT/EP97/01944
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 43804
;   TYPE: DNA
;   ORGANISM: CELO Virus
;   FEATURE:
;   NAME/KEY: gene
;   LOCATION: (12193)..(15043)
;   OTHER INFORMATION: /gene: L1
;   NAME/KEY: misc_feature
;   LOCATION: (15080)
;   OTHER INFORMATION: /note= L2 region penton base splice acceptor site
;   NAME/KEY: gene
;   LOCATION: (15110)..(17495)
;   OTHER INFORMATION: /gene: L2
;   NAME/KEY: polyA_site
;   LOCATION: (17526)
;   NAME/KEY: gene
;   LOCATION: (17559)..(21754)
;   OTHER INFORMATION: /gene: L3
;   NAME/KEY: misc_feature
;   LOCATION: (18261)
;   OTHER INFORMATION: /gene: L3 /note= hexon splice acceptor site
;   NAME/KEY: misc_feature
;   LOCATION: (21102)
;   OTHER INFORMATION: /gene: L3 /note= protease splice acceptor site
;   NAME/KEY: misc_feature
;   LOCATION: (21123)
;   OTHER INFORMATION: /gene: L3 /note= protease splice acceptor site
;   NAME/KEY: polyA_site
;   LOCATION: (21767)
;   NAME/KEY: polyA_site
;   LOCATION: (21824)
;   NAME/KEY: polyA_site
;   LOCATION: (21836)
;   NAME/KEY: polyA_site
;   LOCATION: (21882)
;   NAME/KEY: misc_feature
;   LOCATION: (23608)
;   OTHER INFORMATION: /note= 100K splice acceptor site
;   NAME/KEY: misc_feature
;   LOCATION: (23649)
;   OTHER INFORMATION: /note= 100K splice acceptor site
;   NAME/KEY: gene
;   LOCATION: (23680)..(27886)
;   OTHER INFORMATION: /gene: L4
;   NAME/KEY: polyA_site
;   LOCATION: (27920)
;   NAME/KEY: misc_feature
;   LOCATION: (28315)
;   OTHER INFORMATION: /note= fibre splice acceptor site
;   NAME/KEY: misc_feature
;   LOCATION: (28341)
;   OTHER INFORMATION: / note= fibre splice acceptor site
;   NAME/KEY: gene
;   LOCATION: (28363)..(31768)
;   OTHER INFORMATION: /gene: L5
;   NAME/KEY: misc_feature
;   LOCATION: (30511)
;   OTHER INFORMATION: /gene: L5 /note= fibre splice acceptor site
;   NAME/KEY: polyA_site
;   LOCATION: (31770)
US-09-970-711-1

Alignment Scores:
Pred. No.:              2.44e+03       Length:         43804
Score:                  36.00          Matches:        7
Percent Similarity:     90.00%         Conservative:   2
Best Local Similarity:  70.00%         Mismatches:     1
Query Match:            73.47%         Indels:         0
DB:                     10             Gaps:           0

US-09-823-649A-5 (1-11) x US-09-970-711-1 (1-43804)
```

```
Qy      2 SerValArgLeuGly***ProValLysGlu 11
          ||||||::::::|||   ||||||||||||
Db  32613 TCGGTGCAGGTAGGGGATCCGGTGAAGGAG 32642

RESULT 7
US-10-004-717-18
; Sequence 18, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
;  APPLICANT: ZOGHBI, HUDA Y.
;  APPLICANT:  YANG, QI
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
;  TITLE OF INVENTION:  ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
;  TITLE OF INVENTION:  OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
;  FILE REFERENCE: P01899US4
;  CURRENT APPLICATION NUMBER: US/10/004,717
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: 09/585,645
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: 60/176,993
;  PRIOR FILING DATE: 2000-01-19
;  PRIOR APPLICATION NUMBER: 60/137,060
;  PRIOR FILING DATE: 1999-06-01
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 1074
;   TYPE: DNA
;   ORGANISM: chicken
US-10-004-717-18

Alignment Scores:
Pred. No.:              62          Length:         1074
Score:                  35.00       Matches:        7
Percent Similarity:     80.00%      Conservative:   1
Best Local Similarity:  70.00%      Mismatches:     2
Query Match:            71.43%      Indels:         0
DB:                     9           Gaps:           0

US-09-823-649A-5 (1-11) x US-10-004-717-18 (1-1074)

Qy      2 SerValArgLeuGly***ProValLysGlu 11
          |||:::|||||||||   |||   ||||||
Db    972 TCCATACGACTTGGAAAACCCGGCAAAGAA 1001

RESULT 8
US-09-731-872-37/c
; Sequence 37, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Jobert, Severin
;  TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
;  FILE REFERENCE: 78.US3.REG
;  CURRENT APPLICATION NUMBER: US/09/731,872
;  CURRENT FILING DATE:  2000-12-07
;  PRIOR APPLICATION NUMBER: US 60/169,629
;  PRIOR FILING DATE: 1999-12-08
;  PRIOR APPLICATION NUMBER: US 60/187,470
;  PRIOR FILING DATE: 2000-03-06
;  NUMBER OF SEQ ID NOS: 482
;  SOFTWARE: Patent.pm
; SEQ ID NO 37
;   LENGTH: 1945
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 60..1682
;   NAME/KEY: sig_peptide
;   LOCATION: 60..143
;   OTHER INFORMATION: Von Heijne matrix
;   OTHER INFORMATION: score 3.75144398608723
;   OTHER INFORMATION: seq SGLLLQVLFRLIT/FV
US-09-731-872-37

Alignment Scores:
Pred. No.:              203         Length:         1945
Score:                  34.00       Matches:        7
Percent Similarity:     87.50%      Conservative:   0
Best Local Similarity:  87.50%      Mismatches:     1
Query Match:            69.39%      Indels:         0
DB:                     10          Gaps:           0

US-09-823-649A-5 (1-11) x US-09-731-872-37 (1-1945)

Qy      1 LeuSerValArgLeuGly***Pro 8
          ||||||||||||||||||   |||
Db   1675 TTGTCAGTGCGTCTGGGCACACCT 1652

RESULT 9
US-09-731-872-41/c
; Sequence 41, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Jobert, Severin
;  TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
;  FILE REFERENCE: 78.US3.REG
;  CURRENT APPLICATION NUMBER: US/09/731,872
;  CURRENT FILING DATE:  2000-12-07
;  PRIOR APPLICATION NUMBER: US 60/169,629
;  PRIOR FILING DATE: 1999-12-08
;  PRIOR APPLICATION NUMBER: US 60/187,470
;  PRIOR FILING DATE: 2000-03-06
;  NUMBER OF SEQ ID NOS: 482
;  SOFTWARE: Patent.pm
; SEQ ID NO 41
;   LENGTH: 1953
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 56..1678
;   NAME/KEY: sig_peptide
;   LOCATION: 56..139
;   OTHER INFORMATION: Von Heijne matrix
;   OTHER INFORMATION: score 3.75144398608723
;   OTHER INFORMATION: seq SGLLLQVLFRLIT/FV
US-09-731-872-41

Alignment Scores:
Pred. No.:              204         Length:         1953
Score:                  34.00       Matches:        7
Percent Similarity:     87.50%      Conservative:   0
Best Local Similarity:  87.50%      Mismatches:     1
Query Match:            69.39%      Indels:         0
DB:                     10          Gaps:           0

US-09-823-649A-5 (1-11) x US-09-731-872-41 (1-1953)

Qy      1 LeuSerValArgLeuGly***Pro 8
          ||||||||||||||||||   |||
Db   1671 TTGTCAGTGCGTCTGGGCACACCT 1648

RESULT 10
US-09-731-872-59/c
; Sequence 59, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, Jean Baptiste
;  APPLICANT:  Bougueleret, Lydie
;  APPLICANT:  Jobert, Severin
```

```
;  TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
;  FILE REFERENCE: 78.US3.REG
;  CURRENT APPLICATION NUMBER: US/09/731,872
;  CURRENT FILING DATE:  2000-12-07
;  PRIOR APPLICATION NUMBER: US 60/169,629
;  PRIOR FILING DATE: 1999-12-08
;  PRIOR APPLICATION NUMBER: US 60/187,470
;  PRIOR FILING DATE: 2000-03-06
;  NUMBER OF SEQ ID NOS: 482
;  SOFTWARE: Patent.pm
; SEQ ID NO 59
;   LENGTH: 1969
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 35..1657
;   NAME/KEY: sig_peptide
;   LOCATION: 35..118
;   OTHER INFORMATION: Von Heijne matrix
;   OTHER INFORMATION: score 3.75144398608723
;   OTHER INFORMATION: seq SGLLLQVLFRLIT/FV
US-09-731-872-59

Alignment Scores:
Pred. No.:              205          Length:          1969
Score:                  34.00        Matches:         7
Percent Similarity:     87.50%       Conservative:    0
Best Local Similarity:  87.50%       Mismatches:      1
Query Match:            69.39%       Indels:          0
DB:                     10           Gaps:            0

US-09-823-649A-5 (1-11) x US-09-731-872-59 (1-1969)

Qy       1 LeuSerValArgLeuGly***Pro 8
           ||||||||||||||||||   |||
Db    1650 TTGTCAGTGCGTCTGGGCACACCT 1627

RESULT 11
US-09-815-242-9539
; Sequence 9539, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9539
;   LENGTH: 2001
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2001)
US-09-815-242-9539

Alignment Scores:
Pred. No.:              209          Length:          2001
Score:                  34.00        Matches:         6
Percent Similarity:     81.82%       Conservative:    3
Best Local Similarity:  54.55%       Mismatches:      2
Query Match:            69.39%       Indels:          0
DB:                     10           Gaps:            0

US-09-823-649A-5 (1-11) x US-09-815-242-9539 (1-2001)

Qy       1 LeuSerValArgLeuGly***ProValLysGlu 11
           ||||||||||::::::      |||:::||||||
Db      40 TTATCTGTTAAGGTGAGCTTACCTATAAAGGAA 72

RESULT 12
US-09-815-242-9415
; Sequence 9415, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9415
;   LENGTH: 2031
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2031)
US-09-815-242-9415

Alignment Scores:
Pred. No.:              213          Length:          2031
Score:                  34.00        Matches:         6
Percent Similarity:     81.82%       Conservative:    3
Best Local Similarity:  54.55%       Mismatches:      2
Query Match:            69.39%       Indels:          0
DB:                     10           Gaps:            0
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds
(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSXELSIPYEE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24616/app_query.fasta_1.1393
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649_@CGN_1_1_2820_@runat_21012003_093151_24616 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	90.0	447	13	BI929077	BI929077 EST548966
2	45	90.0	480	12	BF050758	BF050758 EST435916
3	45	90.0	495	13	BI129400	BI129400 G090P24Y
4	45	90.0	507	9	AI895448	AI895448 EST264891
5	45	90.0	534	14	BQ489763	BQ489763 05-E9425-
6	45	90.0	592	10	BE434433	BE434433 EST405511
c 7	45	90.0	629	14	BQ990012	BQ990012 QGF19D10.
8	45	90.0	728	14	BQ856219	BQ856219 QGB28O14.
9	45	90.0	768	13	BM411779	BM411779 EST586106
c 10	41	82.0	788	12	BG216020	BG216020 RST35834
c 11	40	80.0	438	13	BM482565	BM482565 535491 MA
12	39	78.0	183	10	AV415207	AV415207 AV415207
13	39	78.0	411	10	AV428667	AV428667 AV428667
c 14	39	78.0	416	12	BF758509	BF758509 MR4-CT053
15	39	78.0	426	10	AV415722	AV415722 AV415722
16	39	78.0	650	17	AZ826009	AZ826009 2M0101C05
17	39	78.0	658	14	BQ404238	BQ404238 GA__Ed006
c 18	39	78.0	664	14	BQ139058	BQ139058 NF010E06P
c 19	39	78.0	674	13	BI308144	BI308144 EST529554
c 20	39	78.0	742	9	AJ498453	AJ498453 AJ498453
c 21	38	76.0	357	10	AW221400	AW221400 EST297869
c 22	38	76.0	408	12	BF904943	BF904943 H.2-MT017
23	38	76.0	500	17	AQ728992	AQ728992 HS_5466_B
c 24	38	76.0	501	9	AI483959	AI483959 EST249830
c 25	38	76.0	548	10	AW930669	AW930669 EST356512
26	38	76.0	608	13	BM165512	BM165512 EST568035
27	38	76.0	611	17	AQ550437	AQ550437 RPCI-11-3
28	38	76.0	686	17	BH488006	BH488006 BOHJR62TF
29	38	76.0	700	17	BH503133	BH503133 BOGCJ38TF
c 30	37	74.0	246	12	BF553693	BF553693 UI-R-BT0-
c 31	37	74.0	249	9	AV035497	AV035497 AV035497
32	37	74.0	321	14	BQ084906	BQ084906 i145g07.y
33	37	74.0	568	14	BQ696662	BQ696662 NXPV_043_
c 34	37	74.0	571	12	BG834832	BG834832 353186 MA
35	37	74.0	587	13	BJ315731	BJ315731 BJ315731
c 36	37	74.0	652	10	AW319554	AW319554 um97d04.x
37	37	74.0	653	13	BI414990	BI414990 602991072
38	37	74.0	706	13	BI102747	BI102747 602888324
c 39	37	74.0	721	13	BJ321189	BJ321189 BJ321189
40	37	74.0	761	13	BI082481	BI082481 602877828
41	37	74.0	764	9	AU131236	AU131236 AU131236
42	37	74.0	868	9	AU136442	AU136442 AU136442
43	37	74.0	888	13	BI555357	BI555357 603238521
44	37	74.0	928	13	BI903867	BI903867 603166091
c 45	37	74.0	941	12	BG165134	BG165134 602343960

ALIGNMENTS

RESULT 1
BI929077
LOCUS BI929077 447 bp mRNA linear EST 18-OCT-2001
DEFINITION EST548966 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA clone cTOB27L17 5' end, mRNA sequence.
ACCESSION BI929077
VERSION BI929077.1 GI:16241891
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 447)

```
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOB27L17"
                     /clone_lib="tomato flower, 3 - 8 mm buds"
                     /tissue_type="flower"
                     /dev_stage="3-8mm buds"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research. Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT      118 a     92 c     89 g    148 t
ORIGIN

Alignment Scores:
Pred. No.:                52.6          Length:        447
Score:                    45.00         Matches:       8
Percent Similarity:       90.91%        Conservative:  2
Best Local Similarity:    72.73%        Mismatches:    1
Query Match:              90.00%        Indels:        0
DB:                       13            Gaps:          0

US-09-823-649A-4 (1-11) x BI929077 (1-447)

Qy       1 LeuSer***GluLeuSerIleProTyrGluGlu 11
           ||||||   ||||||||||:::||||||||||:::
Db     331 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGGAT 363

RESULT 2
BF050758
LOCUS       BF050758                480 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST435916 tomato developing/immature green fruit Lycopersicon
            esculentum cDNA clone cLEM19O22 5' sequence, mRNA sequence.
ACCESSION   BF050758
VERSION     BF050758.1  GI:10804654
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
            Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
            Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
            ,S.D.
  TITLE     Generation of ESTs from tomato fruit tissue, immature green
  JOURNAL   Unpublished (2000)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEM19O22"
                     /clone_lib="tomato developing/immature green fruit"
                     /tissue_type="fruit"
                     /dev_stage="immature green (5-35 days post-anthesis)"
                     /lab_host="SOLR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoRI;
                     Site_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
                     harvested at 7 day intervals through 35 dpa.  Equal masses
                     of tissue from each stage were combined (including seeds
                     and locules) prior to mRNA isolation."
BASE COUNT      116 a    102 c    119 g    143 t
ORIGIN

Alignment Scores:
Pred. No.:                57.8          Length:        480
Score:                    45.00         Matches:       8
Percent Similarity:       90.91%        Conservative:  2
Best Local Similarity:    72.73%        Mismatches:    1
Query Match:              90.00%        Indels:        0
DB:                       12            Gaps:          0

US-09-823-649A-4 (1-11) x BF050758 (1-480)

Qy       1 LeuSer***GluLeuSerIleProTyrGluGlu 11
           ||||||   ||||||||||:::||||||||||:::
Db     261 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGGAT 293

RESULT 3
BI129400
LOCUS       BI129400                495 bp    mRNA    linear   EST 31-DEC-2001
DEFINITION  G090P24Y Populus cambium cDNA library Populus tremula x Populus
            tremuloides cDNA, mRNA sequence.
ACCESSION   BI129400
VERSION     BI129400.1  GI:18013371
KEYWORDS    EST.
SOURCE      Populus tremula x Populus tremuloides.
  ORGANISM  Populus tremula x Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   Hertzberg,M., Aspeborg,H., Erlandsson,R., Bjorkbacka,H., Hiltonen
            ,T., Karlsson,J., Teeri,T., Gustafsson,P., Bahlerao,R., Jansson,S.,
            Nilsson,O., Sundberg,B., Nilsson,P., Uhlen,M., Sandberg,G. and
            Lundeberg,J.
  TITLE     Gene expression in Populus
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Erlandsson R
            Department of Biotechnology
            Royal Institute of Technology
            Teknikringen 30, Stockholm S-10044, Sweden
            Tel: 46 8 790 8287
            Fax: 46 8 245452
            Email: rikerl@biochem.kth.se.
FEATURES             Location/Qualifiers
     source          1. .495
                     /organism="Populus tremula x Populus tremuloides"
                     /db_xref="taxon:47664"
                     /clone_lib="Populus cambium cDNA library"
                     /note="Organ: cambium"
BASE COUNT      109 a    146 c    117 g    123 t
ORIGIN

Alignment Scores:
Pred. No.:                60.2          Length:        495
Score:                    45.00         Matches:       8
Percent Similarity:       90.91%        Conservative:  2
Best Local Similarity:    72.73%        Mismatches:    1
Query Match:              90.00%        Indels:        0
DB:                       13            Gaps:          0
```

```
US-09-823-649A-4 (1-11) x BI129400 (1-495)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||||||:::||||||||||:::
Db     59 CTATCTGCCGAGTTGTCCGTTCCTTATGAAGAT 91

RESULT 4
AI895448
LOCUS       AI895448                507 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST264891 tomato callus, TAMU Lycopersicon esculentum cDNA clone
            cLEC7H2, mRNA sequence.
ACCESSION   AI895448
VERSION     AI895448.1  GI:5601350
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 507)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
            ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato callus tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence."
FEATURES             Location/Qualifiers
     source          1. .507
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEC7H2"
                     /clone_lib="tomato callus, TAMU"
                     /tissue_type="callus"
                     /dev_stage="25-40 days old"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni laboratory; cLEC - Cotyledons
                     of seedlings 7-10 days post-germination were excised, cut
                     at both ends and placed on MS medium with no selection.
                     Mixed callus was harvested at 25 and 40 days and included
                     undifferentiated masses. Tomato Callus EST Library"
BASE COUNT      125 a    138 c    114 g    130 t
ORIGIN

Alignment Scores:
Pred. No.:                62.1          Length:          507
Score:                    45.00         Matches:         8
Percent Similarity:       90.91%        Conservative:    2
Best Local Similarity:    72.73%        Mismatches:      1
Query Match:              90.00%        Indels:          0
DB:                       9             Gaps:            0

US-09-823-649A-4 (1-11) x AI895448 (1-507)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||||||:::||||||||||:::
Db     66 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGGAT 98

RESULT 5
BQ489763
LOCUS       BQ489763                534 bp    mRNA    linear   EST 07-JUN-2002
DEFINITION  05-E9425-006-009-I01-T3 Sugar beet MPIZ-ADIS-006 Lambda Zap II
            library Beta vulgaris cDNA clone I-1-9, mRNA sequence.
ACCESSION   BQ489763
VERSION     BQ489763.1  GI:21334383
KEYWORDS    EST.
SOURCE      Beta vulgaris.
  ORGANISM  Beta vulgaris
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
REFERENCE   1  (bases 1 to 534)
  AUTHORS   Bellin,D., Werber,M., Theis,T., Weisshaar,B. and Schneider,K.
  TITLE     EST sequencing, annotation and macroarray expression analysis of
            more than 3000 sugar beet cDNAs identifies genes with root-specific
            expression pattern
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Weisshaar B
            ADIS DNA core facility at MPIZ
            Max-Planck-Institute for Plant Breeding Research
            Carl-von-Linne Weg 10, 50829 Koeln, Germany
            Fax: 00492215062851
            Email: weisshaa@mpiz-koeln.mpg.de
            Seq primer: T3 'AATTAACCCTCACTAAAGGG'
            High quality sequence stop: 534.
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Beta vulgaris"
                     /db_xref="taxon:161934"
                     /clone="I-1-9"
                     /clone_lib="Sugar beet MPIZ-ADIS-006 Lambda Zap II
                     library"
                     /dev_stage="4 week old pot-grown plants"
                     /note="Organ: shoot and root; Vector: pBluescript SK- from
                     lambda ZAP II; cDNA (lambda ZAP-II) library from sugar
                     beet, whole plant mRNA, Prepared using the Stratagene
                     UniZAP cDNA kit, cloning sites EcoRI-XhoI, primer sites
                     and orientation:
                     rev-T3-SacI-SK-EcoRI-GGCACGAGG-5pr-cDNA-polyA-XhoI-KpnI-T7
                     -uni"
BASE COUNT      143 a    112 c    127 g    152 t
ORIGIN

Alignment Scores:
Pred. No.:                66.5          Length:          534
Score:                    45.00         Matches:         8
Percent Similarity:       90.91%        Conservative:    2
Best Local Similarity:    72.73%        Mismatches:      1
Query Match:              90.00%        Indels:          0
DB:                       14            Gaps:            0

US-09-823-649A-4 (1-11) x BQ489763 (1-534)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||||||:::||||||||||:::
Db    281 CTCTCTGCTGAACTCTCTGTTCCTTATGAGGAC 313

RESULT 6
BE434433
LOCUS       BE434433                592 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST405511 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
            clone cLEG17C9, mRNA sequence.
ACCESSION   BE434433
VERSION     BE434433.1  GI:9432276
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
            Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
            Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
            ,S.D.
  TITLE     Generation of ESTs from tomato fruit tissue, breaker stage
  JOURNAL   Unpublished (2000)
```

```
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .592
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEG17C9"
                     /clone_lib="tomato breaker fruit, TIGR"
                     /tissue_type="Pericarp"
                     /dev_stage="breaker"
                     /lab_host="SOLR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoR1;
                     Site_2: XhoI; Fruit were harvested at the breaker stage
                     (first sign of lycopene accumulation on the blossom end of
                     the fruit).  Fruit were cut in half and the seeds and
                     locules were discarded prior to freezing the pericarp."
BASE COUNT      140 a    149 c    137 g    166 t
ORIGIN

Alignment Scores:
Pred. No.:              76.1          Length:        592
Score:                  45.00         Matches:       8
Percent Similarity:     90.91%        Conservative:  2
Best Local Similarity:  72.73%        Mismatches:    1
Query Match:            90.00%        Indels:        0
DB:                     10            Gaps:          0

US-09-823-649A-4 (1-11) x BE434433 (1-592)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||||||:::||||||||:::
Db    194 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGGAT 226

RESULT 7
BQ990012/c
LOCUS       BQ990012                629 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  QGF19D10.yg.ab1 QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
            QGF19D10, mRNA sequence.
ACCESSION   BQ990012
VERSION     BQ990012.1  GI:22409537
KEYWORDS    EST.
SOURCE      Lactuca sativa.
  ORGANISM  Lactuca sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE   1  (bases 1 to 629)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Zieqle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QG_CA_Contig1419, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QGF19  row: D  column: 10.
FEATURES             Location/Qualifiers
     source          1. .629
                     /organism="Lactuca sativa"
                     /cultivar="L.serriola"
                     /db_xref="taxon:4236"
                     /clone="QGF19D10"
                     /clone_lib="QG_EFGHJ lettuce serriola"
                     /lab_host="E.coli"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 10 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QG_EFGHJ lettuce serriola
                     TAG_TISSUE=flowers post-fertilized
                     TAG_SEQ=TGCCATCGGC"
BASE COUNT      138 a    137 c    183 g    170 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              82.4          Length:        629
Score:                  45.00         Matches:       8
Percent Similarity:     90.91%        Conservative:  2
Best Local Similarity:  72.73%        Mismatches:    1
Query Match:            90.00%        Indels:        0
DB:                     14            Gaps:          0

US-09-823-649A-4 (1-11) x BQ990012 (1-629)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||||||:::||||||||:::
Db    471 CTTTCAGCAGAACTTTCCGTTCCGTATGAACAT 439

RESULT 8
BQ856219
LOCUS       BQ856219                728 bp    mRNA    linear   EST 14-AUG-2002
DEFINITION  QGB28O14.yg.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone
            QGB28O14, mRNA sequence.
ACCESSION   BQ856219
VERSION     BQ856219.1  GI:22241684
KEYWORDS    EST.
SOURCE      Lactuca sativa.
  ORGANISM  Lactuca sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE   1  (bases 1 to 728)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QG_CA_Contig1419, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QGB28  row: O  column: 14.
FEATURES             Location/Qualifiers
     source          1. .728
                     /organism="Lactuca sativa"
                     /cultivar="Salinas"
                     /db_xref="taxon:4236"
                     /clone="QGB28O14"
                     /clone_lib="QG_ABCDI lettuce salinas"
```

```
                     /lab_host="E.coli"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 10 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QG_ABCDI lettuce salinas
                     TAG_TISSUE=chemical induction
                     TAG_SEQ=TGTAGCCGGG"
BASE COUNT      203 a    203 c    153 g    169 t
ORIGIN

Alignment Scores:
Pred. No.:                99.9          Length:         728
Score:                    45.00         Matches:        8
Percent Similarity:       90.91%        Conservative:   2
Best Local Similarity:    72.73%        Mismatches:     1
Query Match:              90.00%        Indels:         0
DB:                       14            Gaps:           0

US-09-823-649A-4 (1-11) x BQ856219 (1-728)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||||||:::||||||||||:::
Db    299 CTTTCAGCAGAACTTTCCGTTCCGTATGAAGAT 331

RESULT 9
BM411779
LOCUS       BM411779                768 bp    mRNA    linear   EST 22-JAN-2002
DEFINITION  EST586106 tomato breaker fruit Lycopersicon esculentum cDNA clone
            cLEG57L16 5' end, mRNA sequence.
ACCESSION   BM411779
VERSION     BM411779.1  GI:18263409
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai
            ,J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Ronning
            ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue, breaker stage (2002)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1..768
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEG57L16"
                     /clone_lib="tomato breaker fruit"
                     /tissue_type="Pericarp"
                     /dev_stage="breaker"
                     /lab_host="SOLR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoR1;
                     Site_2: XhoI; supplier: Boyce Thompson Institute;
                     sequencing: The Institute for Genomic Research.  Fruit
                     were harvested at the breaker stage (first sign of
                     lycopene accumulation on the blossom end of fruit).  Fruit
                     were cut in half and the seeds and locules were discarded
                     prior to freezing the pericarp."
BASE COUNT      209 a    181 c    173 g    205 t
ORIGIN

Alignment Scores:
Pred. No.:                107           Length:         768
Score:                    45.00         Matches:        8
Percent Similarity:       90.91%        Conservative:   2
Best Local Similarity:    72.73%        Mismatches:     1
Query Match:              90.00%        Indels:         0
DB:                       13            Gaps:           0

US-09-823-649A-4 (1-11) x BM411779 (1-768)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||||||:::||||||||||:::
Db    261 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGGAT 293

RESULT 10
BG216020/c
LOCUS       BG216020                788 bp    mRNA    linear   EST 21-APR-2001
DEFINITION  RST35834 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG216020
VERSION     BG216020.1  GI:13742169
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 788)
  AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
            , Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Oltenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 361 9906
            Fax: 216 361 9599
            Email: scain@athersys.com
            High quality sequence stop: 429.
FEATURES             Location/Qualifiers
     source          1..788
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT      193 a    216 c    198 g    181 t
ORIGIN

Alignment Scores:
Pred. No.:                659           Length:         788
Score:                    41.00         Matches:        8
Percent Similarity:       81.82%        Conservative:   1
Best Local Similarity:    72.73%        Mismatches:     2
Query Match:              82.00%        Indels:         0
DB:                       12            Gaps:           0

US-09-823-649A-4 (1-11) x BG216020 (1-788)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
```

```
          ||||||      ||||||:::||||||||||||
Db    546 CTGTCCCAAAACCTTTCATTGCCTTACGAAGAA 514

RESULT 11
BM482565/c
LOCUS       BM482565 ·                438 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  535491 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BM482565
VERSION     BM482565.1  GI:18532893
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 438)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 3  row: L  column: 18
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .438
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 3BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from marrow, alveolar
                     macrophage, ovary, fetal semitendonosus muscle, and fetal
                     longissimus muscle."
BASE COUNT      126 a     84 c     90 g    138 t
ORIGIN

Alignment Scores:
Pred. No.:              476           Length:        438
Score:                  40.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  1
Best Local Similarity:  87.50%        Mismatches:    0
Query Match:            80.00%        Indels:        0
DB:                     13            Gaps:          0

US-09-823-649A-4 (1-11) x BM482565 (1-438)

Qy      4 GluLeuSerIleProTyrGluGlu 11
          |||||||||:::||||||||||||
Db    313 GAATTATCTGTCCCCTACGAGGAG 290

RESULT 12
AV415207
LOCUS       AV415207                  183 bp    mRNA    linear   EST 23-MAY-2000
DEFINITION  AV415207 Lotus japonicus young plants (two-week old) Lotus
            japonicus cDNA clone MWM104d04_r 5', mRNA sequence.
ACCESSION   AV415207
VERSION     AV415207.1  GI:7744383
KEYWORDS    EST.
SOURCE      Lotus japonicus.
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1  (bases 1 to 183)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
  JOURNAL   DNA Res. 7 (2), 127-130 (2000)
  MEDLINE   20277479
COMMENT     Contact: Yasukazu Nakamura
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .183
                     /organism="Lotus japonicus"
                     /db_xref="taxon:34305"
                     /clone="MWM104d04_r"
                     /clone_lib="Lotus japonicus young plants (two-week old)"
                     /dev_stage="young plants (two-week old)"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; isolate=Miyakojima MG-20"
BASE COUNT       33 a     64 c     44 g     42 t
ORIGIN

Alignment Scores:
Pred. No.:              236           Length:        183
Score:                  39.00         Matches:       7
Percent Similarity:     81.82%        Conservative:  2
Best Local Similarity:  63.64%        Mismatches:    2
Query Match:            78.00%        Indels:        0
DB:                     10            Gaps:          0

US-09-823-649A-4 (1-11) x AV415207 (1-183)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
            ||||||   ||||||   :::|||||||:::|||
Db    117 CTCTCCGGCGAGCTCTGTGTTCCGTATGATGAA 149

RESULT 13
AV428667
LOCUS       AV428667                  411 bp    mRNA    linear   EST 02-MAY-2000
DEFINITION  AV428667 Lotus japonicus young plants (two-week old) Lotus
            japonicus cDNA clone MWL050q03_r 5', mRNA sequence.
ACCESSION   AV428667
VERSION     AV428667.1  GI:7678049
KEYWORDS    EST.
SOURCE      Lotus japonicus.
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
  JOURNAL   DNA Res. 7 (2), 127-130 (2000)
  MEDLINE   20277479
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .411
                     /organism="Lotus japonicus"
```

```
                     /db_xref="taxon:34305"
                     /clone="MWL050g03_r"
                     /clone_lib="Lotus japonicus young plants (two-week old)"
                     /dev_stage="young plants (two-week old)"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; isolate=Miyakojima MG-20"
BASE COUNT       86 a    143 c     90 g     90 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              683           Length:        411
Score:                  39.00         Matches:       7
Percent Similarity:     81.82%        Conservative:  2
Best Local Similarity:  63.64%        Mismatches:    2
Query Match:            78.00%        Indels:        0
DB:                     10            Gaps:          0

US-09-823-649A-4 (1-11) x AV428667 (1-411)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||   :::|||||||:::|||
Db    374 CTCTCCGGCGAGCTCTGTGTNCCGTATGATGAA 406

RESULT 14
BF758509/c
LOCUS       BF758509                416 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  MR4-CT0538-141100-104-b08 CT0538 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF758509
VERSION     BF758509.1  GI:12106409
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-CT0538-
            141100-104-b08&t3=2000-11-14&t4-1)
            Seq primer: puc 18 forward
            High quality sequence start: 30
            High quality sequence stop: 88.
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="CT0538"
                     /dev_stage="Adult"
                     /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT      129 a    106 c     85 g     96 t
ORIGIN

Alignment Scores:
Pred. No.:              694           Length:        416
Score:                  39.00         Matches:       8
Percent Similarity:     81.82%        Conservative:  1
Best Local Similarity:  72.73%        Mismatches:    2
Query Match:            78.00%        Indels:        0
DB:                     12            Gaps:          0

US-09-823-649A-4 (1-11) x BF758509 (1-416)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   |||   :::|||||||||||||||
Db    182 TTGTCTGGTGAGCGAGCTATCCCTTATGAGGAG 150

RESULT 15
AV415722
LOCUS       AV415722                426 bp    mRNA    linear   EST 23-MAY-2000
DEFINITION  AV415722 Lotus japonicus young plants (two-week old) Lotus
            japonicus cDNA clone MWM115e12_r 5', mRNA sequence.
ACCESSION   AV415722
VERSION     AV415722.1  GI:7744898
KEYWORDS    EST.
SOURCE      Lotus japonicus.
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
  JOURNAL   DNA Res. 7 (2), 127-130 (2000)
  MEDLINE   20277479
COMMENT     Contact: Yasukazu Nakamura
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Lotus japonicus"
                     /db_xref="taxon:34305"
                     /clone="MWM115e12_r"
                     /clone_lib="Lotus japonicus young plants (two-week old)"
                     /dev_stage="young plants (two-week old)"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; isolate=Miyakojima MG-20"
BASE COUNT       90 a    147 c     95 g     94 t
ORIGIN

Alignment Scores:
Pred. No.:              716           Length:        426
Score:                  39.00         Matches:       7
Percent Similarity:     81.82%        Conservative:  2
Best Local Similarity:  63.64%        Mismatches:    2
Query Match:            78.00%        Indels:        0
DB:                     10            Gaps:          0

US-09-823-649A-4 (1-11) x AV415722 (1-426)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||   :::|||||||:::|||
Db    375 CTCTCCGGCGAGCTCTGTGTTCCGTATGATGAA 407


Search completed: January 21, 2003, 12:24:18
Job time : 1170.14 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds
(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSXELSIPYEE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: ..
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24626/app_query.fasta_1.1393
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human4[illegible]cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649_@CGN_1_1_77_@runat_21012003_093151_24626
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	96.0	2496	9	US-10-033-297-2	Sequence 2, Appli
2	48	96.0	2496	9	US-10-081-806-2	Sequence 2, Appli
3	45	90.0	1600	9	US-10-033-297-12	Sequence 12, Appl
4	45	90.0	1600	9	US-10-081-806-12	Sequence 12, Appl
5	45	90.0	2502	9	US-10-033-297-7	Sequence 7, Appli
6	45	90.0	2502	9	US-10-033-297-21	Sequence 21, Appl
7	45	90.0	2502	9	US-10-081-806-7	Sequence 7, Appli
8	45	90.0	2502	9	US-10-081-806-21	Sequence 21, Appl
9	45	90.0	2504	9	US-10-033-297-3	Sequence 3, Appli
10	45	90.0	2504	9	US-10-081-806-3	Sequence 3, Appli
11	45	90.0	2505	9	US-10-033-297-65	Sequence 65, Appl
12	45	90.0	2505	9	US-10-033-297-68	Sequence 68, Appl
13	45	90.0	2505	9	US-10-033-297-70	Sequence 70, Appl
14	45	90.0	2506	9	US-10-033-297-1	Sequence 1, Appli
15	45	90.0	2506	9	US-10-081-806-1	Sequence 1, Appli
16	45	90.0	2511	10	US-09-777-430A-9	Sequence 9, Appli
17	45	90.0	2511	10	US-09-777-430A-12	Sequence 12, Appl
18	45	90.0	2526	10	US-09-777-430A-14	Sequence 14, Appl
19	45	90.0	2526	10	US-09-777-430A-18	Sequence 19, Appl
20	45	90.0	2526	10	US-09-777-430A-22	Sequence 22, Appl
21	45	90.0	2526	10	US-09-777-430A-25	Sequence 25, Appl
22	45	90.0	2626	9	US-09-972-834-1	Sequence 1, Appli
23	37	74.0	1497	10	US-09-735-169A-3	Sequence 3, Appli
24	37	74.0	1497	10	US-09-735-171A-3	Sequence 3, Appli
25	37	74.0	3452	10	US-09-735-169A-1	Sequence 1, Appli
26	37	74.0	3452	10	US-09-735-171A-1	Sequence 1, Appli
c 27	36	72.0	286	10	US-09-923-876-6139	Sequence 6139, Ap
28	36	72.0	2603	9	US-09-764-868-80	Sequence 80, Appl
c 29	34	68.0	302	9	US-10-046-935-1913	Sequence 1913, Ap
c 30	34	68.0	302	9	US-09-878-178-1913	Sequence 1913, Ap
31	34	68.0	378	10	US-09-734-017A-57	Sequence 57, Appl
32	34	68.0	402	10	US-09-974-300-3201	Sequence 3201, Ap
c 33	34	68.0	675	10	US-09-841-132-377	Sequence 377, App
c 34	34	68.0	675	10	US-09-841-132-473	Sequence 473, App
35	34	68.0	1399	10	US-09-764-870-134	Sequence 134, App
c 36	34	68.0	1812	10	US-09-925-300-724	Sequence 724, App
c 37	34	68.0	2068	12	US-10-044-090-519	Sequence 519, App
c 38	34	68.0	2515	12	US-10-044-090-520	Sequence 520, App
39	34	68.0	3664	10	US-09-764-877-2933	Sequence 2933, Ap
c 40	34	68.0	12445	10	US-09-070-927A-242	Sequence 242, App
41	34	68.0	15295	10	US-09-764-877-3404	Sequence 3404, Ap
42	34	68.0	16281	10	US-09-764-847-1367	Sequence 1367, Ap
43	34	68.0	16285	10	US-09-764-847-1368	Sequence 1368, Ap
44	34	68.0	16285	10	US-09-764-847-1369	Sequence 1369, Ap
c 45	34	68.0	30515	10	US-09-764-847-1208	Sequence 1208, Ap

ALIGNMENTS

RESULT 1
US-10-033-297-2
; Sequence 2, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; Lyamichev, Victor I.
; Mast, Andrea L.
; Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/033,297
; FILING DATE: 12-No. US20020187486A1-2001
; CLASSIFICATION: <Unknown>

```
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
;            FILING DATE: 21-JAN-1997
;            APPLICATION NUMBER: US 08/759,038
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/758,314
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/756,386
;            FILING DATE: 29-NOV-1996
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2496 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-033-297-2

Alignment Scores:
Pred. No.:              0.501          Length:          2496
Score:                  48.00          Matches:         10
Percent Similarity:     90.91%         Conservative:    0
Best Local Similarity:  90.91%         Mismatches:      1
Query Match:            96.00%         Indels:          0
DB:                     9              Gaps:            0

US-09-823-649A-4 (1-11) x US-10-033-297-2 (1-2496)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||||||||||||||||||||
Db   2029 CTCTCCGGGGAGCTTTCCATCCCCTACGAGGAG 2061

RESULT 2
US-10-081-806-2
; Sequence 2, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;       APPLICANT: Prudent, James R.
;                  Hall, Jeff G.
;                  Lyamichev, Victor I.
;       TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;       NUMBER OF SEQUENCES: 69
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/081,806
;            FILING DATE: 22-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/756,386
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02564
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2496 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-081-806-2

Alignment Scores:
Pred. No.:              0.501          Length:          2496
Score:                  48.00          Matches:         10
Percent Similarity:     90.91%         Conservative:    0
Best Local Similarity:  90.91%         Mismatches:      1
Query Match:            96.00%         Indels:          0
DB:                     9              Gaps:            0

US-09-823-649A-4 (1-11) x US-10-081-806-2 (1-2496)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||||||||||||||||||||
Db   2029 CTCTCCGGGGAGCTTTCCATCCCCTACGAGGAG 2061

RESULT 3
US-10-033-297-12
; Sequence 12, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/033,297
;            FILING DATE: 12-No. US20020187486A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
```

```
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 1600 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-033-297-12

Alignment Scores:
Pred. No.:                1.32          Length:          1600
Score:                    45.00         Matches:         9
Percent Similarity:       90.91%        Conservative:    1
Best Local Similarity:    81.82%        Mismatches:      1
Query Match:              90.00%        Indels:          0
DB:                       9             Gaps:            0

US-09-823-649A-4 (1-11) x US-10-033-297-12 (1-1600)

Qy       1 LeuSer***GluLeuSerIleProTyrGluGlu 11
           ||||||   ||||||:::|||||||||||||||||
Db    1133 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 1165

RESULT 4
US-10-081-806-12
; Sequence 12, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;       APPLICANT: Prudent, James R.
;                  Hall, Jeff G.
;                  Lyamichev, Victor I.
;       TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;       NUMBER OF SEQUENCES: 69
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 1600 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-081-806-12

Alignment Scores:
Pred. No.:                1.32          Length:          1600
Score:                    45.00         Matches:         9
Percent Similarity:       90.91%        Conservative:    1
Best Local Similarity:    81.82%        Mismatches:      1
Query Match:              90.00%        Indels:          0
DB:                       9             Gaps:            0

US-09-823-649A-4 (1-11) x US-10-081-806-12 (1-1600)

Qy       1 LeuSer***GluLeuSerIleProTyrGluGlu 11
           ||||||   ||||||:::|||||||||||||||||
Db    1133 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 1165

RESULT 5
US-10-033-297-7
; Sequence 7, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
```

```
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-033-297-7

Alignment Scores:
Pred. No.:              2.19          Length:          2502
Score:                  45.00         Matches:         9
Percent Similarity:     90.91%        Conservative:    1
Best Local Similarity:  81.82%        Mismatches:      1
Query Match:            90.00%        Indels:          0
DB:                     9             Gaps:            0

US-09-823-649A-4 (1-11) x US-10-033-297-7 (1-2502)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||:::|||||||||||||||||
Db   2035 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2067

RESULT 6
US-10-033-297-21
; Sequence 21, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 21:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-033-297-21

Alignment Scores:
Pred. No.:              2.19          Length:          2502
Score:                  45.00         Matches:         9
Percent Similarity:     90.91%        Conservative:    1
Best Local Similarity:  81.82%        Mismatches:      1
Query Match:            90.00%        Indels:          0
DB:                     9             Gaps:            0

US-09-823-649A-4 (1-11) x US-10-033-297-21 (1-2502)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||:::|||||||||||||||||
Db   2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 7
US-10-081-806-7
; Sequence 7, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;       APPLICANT: Prudent, James R.
;                  Hall, Jeff G.
;                  Lyamichev, Victor I.
;       TITLE OF INVENTION: Invasive Cleavage of Nucleic Acids
;       NUMBER OF SEQUENCES: 69
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
```

```
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-081-806-7

Alignment Scores:
Pred. No.:                2.19          Length:          2502
Score:                    45.00         Matches:         9
Percent Similarity:       90.91%        Conservative:    1
Best Local Similarity:    81.82%        Mismatches:      1
Query Match:              90.00%        Indels:          0
DB:                       9             Gaps:            0

US-09-823-649A-4 (1-11) x US-10-081-806-7 (1-2502)

Qy       1 LeuSer***GluLeuSerIleProTyrGluGlu 11
           ||||||   ||||||:::|||||||||||||||||
Db    2035 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2067

RESULT 8
US-10-081-806-21
; Sequence 21, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;        APPLICANT: Prudent, James R.
;                   Hall, Jeff G.
;                   Lyamichev, Victor I.
;        TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;        NUMBER OF SEQUENCES: 69
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 21:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-081-806-21

Alignment Scores:
Pred. No.:                2.19          Length:          2502
Score:                    45.00         Matches:         9
Percent Similarity:       90.91%        Conservative:    1
Best Local Similarity:    81.82%        Mismatches:      1
Query Match:              90.00%        Indels:          0
DB:                       9             Gaps:            0

US-09-823-649A-4 (1-11) x US-10-081-806-21 (1-2502)

Qy       1 LeuSer***GluLeuSerIleProTyrGluGlu 11
           ||||||   ||||||:::|||||||||||||||||
Db    2035 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2067

RESULT 9
US-10-033-297-3
; Sequence 3, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-Nov-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
```

```
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2504 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-033-297-3

Alignment Scores:
Pred. No.:              2.19          Length:         2504
Score:                  45.00         Matches:        9
Percent Similarity:     90.91%        Conservative:   1
Best Local Similarity:  81.82%        Mismatches:     1
Query Match:            90.00%        Indels:         0
DB:                     9             Gaps:           0

US-09-823-649A-4 (1-11) x US-10-033-297-3 (1-2504)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||:::|||||||||||||||
Db   2038 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2070

RESULT 10
US-10-081-806-3
; Sequence 3, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;        APPLICANT: Prudent, James R.
;                   Hall, Jeff G.
;                   Lyamichev, Victor I.
;        TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;        NUMBER OF SEQUENCES: 69
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2504 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-081-806-3

Alignment Scores:
Pred. No.:              2.19          Length:         2504
Score:                  45.00         Matches:        9
Percent Similarity:     90.91%        Conservative:   1
Best Local Similarity:  81.82%        Mismatches:     1
Query Match:            90.00%        Indels:         0
DB:                     9             Gaps:           0

US-09-823-649A-4 (1-11) x US-10-081-806-3 (1-2504)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||:::|||||||||||||||
Db   2038 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2070

RESULT 11
US-10-033-297-65
; Sequence 65, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 65:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2505 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
```

```
;       MOLECULE TYPE: DNA (genomic)
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..2499
;       SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-10-033-297-65

Alignment Scores:
Pred. No.:              2.19          Length:         2505
Score:                  45.00         Matches:        9
Percent Similarity:     90.91%        Conservative:   1
Best Local Similarity:  81.82%        Mismatches:     1
Query Match:            90.00%        Indels:         0
DB:                     9             Gaps:           0

US-09-823-649A-4 (1-11) x US-10-033-297-65 (1-2505)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||:::|||||||||||||||||
Db   2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 12
US-10-033-297-68
; Sequence 68, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/033,297
;            FILING DATE: 12-No. US20020187486A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
;            FILING DATE: 21-JAN-1997
;            APPLICATION NUMBER: US 08/759,038
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/758,314
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/756,386
;            FILING DATE: 29-NOV-1996
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
```

```
;   INFORMATION FOR SEQ ID NO: 68:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2505 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: other nucleic acid
;            DESCRIPTION: /desc = "DNA"
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..2499
;       SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-10-033-297-68

Alignment Scores:
Pred. No.:              2.19          Length:         2505
Score:                  45.00         Matches:        9
Percent Similarity:     90.91%        Conservative:   1
Best Local Similarity:  81.82%        Mismatches:     1
Query Match:            90.00%        Indels:         0
DB:                     9             Gaps:           0

US-09-823-649A-4 (1-11) x US-10-033-297-68 (1-2505)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||:::|||||||||||||||||
Db   2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 13
US-10-033-297-70
; Sequence 70, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/033,297
;            FILING DATE: 12-No. US20020187486A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
;            FILING DATE: 21-JAN-1997
;            APPLICATION NUMBER: US 08/759,038
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/758,314
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/756,386
;            FILING DATE: 29-NOV-1996
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
```

```
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 70:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2505 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: other nucleic acid
;            DESCRIPTION: /desc = "DNA"
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..2499
;       SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-10-033-297-70

Alignment Scores:
Pred. No.:              2.19          Length:         2505
Score:                  45.00         Matches:        9
Percent Similarity:     90.91%        Conservative:   1
Best Local Similarity:  81.82%        Mismatches:     1
Query Match:            90.00%        Indels:         0
DB:                     9             Gaps:           0

US-09-823-649A-4 (1-11) x US-10-033-297-70 (1-2505)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||:::|||||||||||||||
Db   2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 14
US-10-033-297-1
; Sequence 1, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/033,297
;            FILING DATE: 12-No. US20020187486A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
;            FILING DATE: 21-JAN-1997
;            APPLICATION NUMBER: US 08/759,038
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/758,314
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/756,386
;            FILING DATE: 29-NOV-1996
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2506 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-033-297-1

Alignment Scores:
Pred. No.:              2.19          Length:         2506
Score:                  45.00         Matches:        9
Percent Similarity:     90.91%        Conservative:   1
Best Local Similarity:  81.82%        Mismatches:     1
Query Match:            90.00%        Indels:         0
DB:                     9             Gaps:           0

US-09-823-649A-4 (1-11) x US-10-033-297-1 (1-2506)

Qy      1 LeuSer***GluLeuSerIleProTyrGluGlu 11
          ||||||   ||||||:::|||||||||||||||
Db   2032 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2064

RESULT 15
US-10-081-806-1
; Sequence 1, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;       APPLICANT: Prudent, James R.
;                  Hall, Jeff G.
;                  Lyamichev, Victor I.
;       TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;       NUMBER OF SEQUENCES: 69
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/081,806
;            FILING DATE: 22-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/756,386
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
```

```
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2506 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-081-806-1

Alignment Scores:
Pred. No.:              2.19          Length:          2506
Score:                  45.00         Matches:         9
Percent Similarity:     90.91%        Conservative:    1
Best Local Similarity:  81.82%        Mismatches:      1
Query Match:            90.00%        Indels:          0
DB:                     9             Gaps:            0

US-09-823-649A-4 (1-11) x US-10-081-806-1 (1-2506)

Qy       1 LeuSer***GluLeuSerIleProTyrGluGlu 11
           ||||||   ||||||:::|||||||||||||||
Db    2032 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2064


Search completed: January 21, 2003, 10:07:46
Job time : 31.5714 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds
(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649A-3
Perfect score: 54
Sequence: 1 LSQELAIPYEE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24626/app_query.fasta_1.1393
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649_@CGN_1_1_77_@runat_21012003_093151_24626
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	54	100.0	1600	9	US-10-033-297-12	Sequence 12, Appl
2	54	100.0	1600	9	US-10-081-806-12	Sequence 12, Appl
3	54	100.0	2502	9	US-10-033-297-7	Sequence 7, Appli
4	54	100.0	2502	9	US-10-033-297-21	Sequence 21, Appl
5	54	100.0	2502	9	US-10-081-806-7	Sequence 7, Appli
6	54	100.0	2502	9	US-10-081-806-21	Sequence 21, Appl
7	54	100.0	2504	9	US-10-033-297-3	Sequence 3, Appli
8	54	100.0	2504	9	US-10-081-806-3	Sequence 3, Appli
9	54	100.0	2505	9	US-10-033-297-65	Sequence 65, Appl
10	54	100.0	2505	9	US-10-033-297-68	Sequence 68, Appl
11	54	100.0	2505	9	US-10-033-297-70	Sequence 70, Appl
12	54	100.0	2506	9	US-10-033-297-1	Sequence 1, Appli
13	54	100.0	2506	9	US-10-081-806-1	Sequence 1, Appli
14	54	100.0	2511	10	US-09-777-430A-9	Sequence 9, Appli
15	54	100.0	2511	10	US-09-777-430A-12	Sequence 12, Appl
16	54	100.0	2526	10	US-09-777-430A-14	Sequence 14, Appl
17	54	100.0	2526	10	US-09-777-430A-19	Sequence 19, Appl
18	54	100.0	2526	10	US-09-777-430A-22	Sequence 22, Appl
19	54	100.0	2526	10	US-09-777-430A-25	Sequence 25, Appl
20	54	100.0	2626	9	US-09-972-834-1	Sequence 1, Appli
21	44	81.5	2496	9	US-10-033-297-2	Sequence 2, Appli
22	44	81.5	2496	9	US-10-081-806-2	Sequence 2, Appli
c 23	40	74.1	24	9	US-09-972-834-9	Sequence 9, Appli
24	39	72.2	2603	9	US-09-764-868-80	Sequence 80, Appl
c 25	39	72.2	32768	10	US-09-070-927A-128	Sequence 128, App
c 26	38	70.4	286	10	US-09-923-876-6139	Sequence 6139, Ap
27	37	68.5	481	10	US-09-920-300A-818	Sequence 818, App
28	37	68.5	481	12	US-10-033-528-818	Sequence 818, App
c 29	37	68.5	1812	10	US-09-925-300-724	Sequence 724, App
30	36	66.7	319	10	US-09-881-752A-301	Sequence 301, App
31	36	66.7	719	9	US-09-895-913A-141	Sequence 141, App
32	36	66.7	819	10	US-09-815-242-6393	Sequence 6393, Ap
33	36	66.7	2088	9	US-10-033-297-10	Sequence 10, Appl
34	36	66.7	2088	9	US-10-081-806-10	Sequence 10, Appl
35	36	66.7	10397	10	US-09-070-927A-290	Sequence 290, App
c 36	35	64.8	403	10	US-09-815-242-1034	Sequence 1034, Ap
c 37	35	64.8	410	10	US-09-960-352-1424	Sequence 1424, Ap
c 38	35	64.8	1044	10	US-09-816-127-4	Sequence 4, Appli
39	35	64.8	1755	10	US-09-815-242-3918	Sequence 3918, Ap
40	35	64.8	1761	10	US-09-815-242-6483	Sequence 6483, Ap
41	35	64.8	1828	9	US-10-114-893-53	Sequence 53, Appl
42	35	64.8	3065	10	US-09-764-877-3991	Sequence 3991, Ap
43	35	64.8	3465	10	US-09-739-254-49	Sequence 49, Appl
44	35	64.8	3465	10	US-09-904-615-49	Sequence 49, Appl
c 45	35	64.8	7893	10	US-09-070-927A-228	Sequence 228, App

ALIGNMENTS

```
RESULT 1
US-10-033-297-12
; Sequence 12, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
```

```
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
;            FILING DATE: 21-JAN-1997
;            APPLICATION NUMBER: US 08/759,038
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/758,314
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/756,386
;            FILING DATE: 29-NOV-1996
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1600 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-033-297-12

Alignment Scores:
Pred. No.:                0.0165          Length:          1600
Score:                    54.00           Matches:         11
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-823-649A-3 (1-11) x US-10-033-297-12 (1-1600)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    1133 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 1165

RESULT 2
US-10-081-806-12
; Sequence 12, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;       APPLICANT: Prudent, James R.
;                  Hall, Jeff G.
;                  Lyamichev, Victor I.
;       TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;       NUMBER OF SEQUENCES: 69
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/081,806
;            FILING DATE: 22-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/756,386
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02564
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1600 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-081-806-12

Alignment Scores:
Pred. No.:                0.0165          Length:          1600
Score:                    54.00           Matches:         11
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-823-649A-3 (1-11) x US-10-081-806-12 (1-1600)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    1133 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 1165

RESULT 3
US-10-033-297-7
; Sequence 7, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/033,297
;            FILING DATE: 12-No. US20020187486A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
```

```
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-033-297-7

Alignment Scores:
Pred. No.:                0.0276          Length:          2502
Score:                    54.00           Matches:         11
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-823-649A-3 (1-11) x US-10-033-297-7 (1-2502)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    2035 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2067

RESULT 4
US-10-033-297-21
; Sequence 21, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 21:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-033-297-21

Alignment Scores:
Pred. No.:                0.0276          Length:          2502
Score:                    54.00           Matches:         11
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-823-649A-3 (1-11) x US-10-033-297-21 (1-2502)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 5
US-10-081-806-7
; Sequence 7, Application US/10081806
; Publication No. US20020197624A1
;   GENERAL INFORMATION:
;       APPLICANT: Prudent, James R.
;                  Hall, Jeff G.
;                  Lyamichev, Victor I.
;       TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;       NUMBER OF SEQUENCES: 69
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
```

```
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-081-806-7

Alignment Scores:
Pred. No.:              0.0276          Length:          2502
Score:                  54.00           Matches:         11
Percent Similarity:     100.00%         Conservative:    0
Best Local Similarity:  100.00%         Mismatches:      0
Query Match:            100.00%         Indels:          0
DB:                     9               Gaps:            0

US-09-823-649A-3 (1-11) x US-10-081-806-7 (1-2502)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    2035 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2067

RESULT 6
US-10-081-806-21
; Sequence 21, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;        APPLICANT: Prudent, James R.
;                   Hall, Jeff G.
;                   Lyamichev, Victor I.
;        TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;        NUMBER OF SEQUENCES: 69
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 21:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-081-806-21

Alignment Scores:
Pred. No.:              0.0276          Length:          2502
Score:                  54.00           Matches:         11
Percent Similarity:     100.00%         Conservative:    0
Best Local Similarity:  100.00%         Mismatches:      0
Query Match:            100.00%         Indels:          0
DB:                     9               Gaps:            0

US-09-823-649A-3 (1-11) x US-10-081-806-21 (1-2502)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 7
US-10-033-297-3
; Sequence 3, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
```

```
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2504 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-033-297-3

Alignment Scores:
Pred. No.:                0.0277          Length:          2504
Score:                    54.00           Matches:         11
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-823-649A-3 (1-11) x US-10-033-297-3 (1-2504)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    2038 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2070

RESULT 8
US-10-081-806-3
; Sequence 3, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;       APPLICANT: Prudent, James R.
;                  Hall, Jeff G.
;                  Lyamichev, Victor I.
;       TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;       NUMBER OF SEQUENCES: 69
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2504 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-081-806-3

Alignment Scores:
Pred. No.:                0.0277          Length:          2504
Score:                    54.00           Matches:         11
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-823-649A-3 (1-11) x US-10-081-806-3 (1-2504)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    2038 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2070

RESULT 9
US-10-033-297-65
; Sequence 65, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
```

```
;   INFORMATION FOR SEQ ID NO: 65:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2505 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..2499
;        SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-10-033-297-65

Alignment Scores:
Pred. No.:              0.0277          Length:         2505
Score:                  54.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-823-649A-3 (1-11) x US-10-033-297-65 (1-2505)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||||||||||||||||||||||||||
Db   2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 10
US-10-033-297-68
; Sequence 68, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 68:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2505 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: other nucleic acid
;             DESCRIPTION: /desc = "DNA"
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..2499
;        SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-10-033-297-68

Alignment Scores:
Pred. No.:              0.0277          Length:         2505
Score:                  54.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-823-649A-3 (1-11) x US-10-033-297-68 (1-2505)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||||||||||||||||||||||||||
Db   2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 11
US-10-033-297-70
; Sequence 70, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
```

```
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 70:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2505 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: other nucleic acid
;             DESCRIPTION: /desc = "DNA"
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  1..2499
;        SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-10-033-297-70

Alignment Scores:
Pred. No.:              0.0277          Length:         2505
Score:                  54.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-823-649A-3 (1-11) x US-10-033-297-70 (1-2505)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 12
US-10-033-297-1
; Sequence 1, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2506 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-033-297-1

Alignment Scores:
Pred. No.:              0.0277          Length:         2506
Score:                  54.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-823-649A-3 (1-11) x US-10-033-297-1 (1-2506)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||||||||||||||||||||||||
Db    2032 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2064

RESULT 13
US-10-081-806-1
; Sequence 1, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;        APPLICANT: Prudent, James R.
;                   Hall, Jeff G.
;                   Lyamichev, Victor I.
;        TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
;        NUMBER OF SEQUENCES: 69
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
```

```
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2506 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-081-806-1

Alignment Scores:
Pred. No.:              0.0277          Length:         2506
Score:                  54.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-823-649A-3 (1-11) x US-10-081-806-1 (1-2506)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||||||||||||||||||||||||||
Db   2032 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2064

RESULT 14
US-09-777-430A-9
; Sequence 9, Application US/09777430A
; Patent No. US20020128465A1
; GENERAL INFORMATION:
;  APPLICANT: Lyamichev, Victor
;  APPLICANT:  Skrzypczynski, Zbigniew
;  APPLICANT:  Allawi, Hatim T.
;  APPLICANT:  Wayland, Sarah R.
;  APPLICANT:  Takova, Tsetska
;  APPLICANT:  Neir, Bruce P.
;  TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
;  FILE REFERENCE: FORS-04912
;  CURRENT APPLICATION NUMBER: US/09/777,430A
;  CURRENT FILING DATE:  2001-02-06
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 2511
;   TYPE: DNA
;   ORGANISM: Thermus thermophilus
US-09-777-430A-9

Alignment Scores:
Pred. No.:              0.0278          Length:         2511
Score:                  54.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     10              Gaps:           0

US-09-823-649A-3 (1-11) x US-09-777-430A-9 (1-2511)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||||||||||||||||||||||||||
Db   2044 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2076

RESULT 15
US-09-777-430A-12
; Sequence 12, Application US/09777430A
; Patent No. US20020128465A1
; GENERAL INFORMATION:
;  APPLICANT: Lyamichev, Victor
;  APPLICANT:  Skrzypczynski, Zbigniew
;  APPLICANT:  Allawi, Hatim T.
;  APPLICANT:  Wayland, Sarah R.
;  APPLICANT:  Takova, Tsetska
;  APPLICANT:  Neir, Bruce P.
;  TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
;  FILE REFERENCE: FORS-04912
;  CURRENT APPLICATION NUMBER: US/09/777,430A
;  CURRENT FILING DATE:  2001-02-06
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 2511
;   TYPE: DNA
;   ORGANISM: Thermus thermophilus
US-09-777-430A-12

Alignment Scores:
Pred. No.:              0.0278          Length:         2511
Score:                  54.00           Matches:        11
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     10              Gaps:           0

US-09-823-649A-3 (1-11) x US-09-777-430A-12 (1-2511)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||||||||||||||||||||||||||
Db   2044 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2076


Search completed: January 21, 2003, 10:07:44
Job time : 32.5714 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds
(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649A-3
Perfect score: 54
Sequence: 1 LSQELAIPYEE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24616/app_query.fasta_1.1393
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649_@CGN_1_1_2820_@runat_21012003_093151_24616 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	44	81.5	788	12	BG216020	BG216020 RST35834
c 2	42	77.8	319	13	BI921437	BI921437 EST541340
c 3	42	77.8	467	12	BF050176	BF050176 EST435334
c 4	42	77.8	526	10	AW932704	AW932704 EST358547
c 5	42	77.8	528	13	BI932551	BI932551 EST552440
c 6	42	77.8	529	10	AW031837	AW031837 EST275291
c 7	42	77.8	537	10	AW399580	AW399580 EST310080
c 8	42	77.8	543	10	AW930945	AW930945 EST356788
c 9	42	77.8	558	10	AW931411	AW931411 EST357254
c 10	42	77.8	590	10	AW931055	AW931055 EST356898
c 11	42	77.8	622	10	AW737720	AW737720 EST339147
12	42	77.8	653	10	AW036109	AW036109 EST274485
c 13	42	77.8	743	13	BI933283	BI933283 EST553172
c 14	42	77.8	759	12	BG124823	BG124823 EST470469
15	41	75.9	198	9	AV043104	AV043104 AV043104
c 16	41	75.9	416	12	BF758509	BF758509 MR4-CT053
17	41	75.9	447	13	BI929077	BI929077 EST548966
18	41	75.9	480	12	BF050758	BF050758 EST435916
19	41	75.9	495	13	BI129400	BI129400 G090P24Y
20	41	75.9	507	9	AI895448	AI895448 EST264891
21	41	75.9	534	14	BQ489763	BQ489763 05-E9425-
22	41	75.9	592	10	BE434433	BE434433 EST405511
c 23	41	75.9	629	14	BQ990012	BQ990012 QGF19D10.
24	41	75.9	712	17	BH702922	BH702922 BOHXK45TF
25	41	75.9	728	14	BQ856219	BQ856219 QGB28014.
26	41	75.9	768	13	BM411779	BM411779 EST586106
27	41	75.9	830	17	BH681916	BH681916 BOMNR23TF
c 28	40	74.1	249	9	AV035497	AV035497 AV035497
29	40	74.1	398	9	AI356938	AI356938 qx21d11.x
c 30	40	74.1	451	10	AW644700	AW644700 cm44b06.w
c 31	40	74.1	491	12	BF325103	BF325103 su30g06.y
c 32	40	74.1	583	12	BG459033	BG459033 00806 lea
c 33	40	74.1	752	17	BH353192	BH353192 CH230-221
34	40	74.1	758	17	AZ191827	AZ191827 SP_1020_A
35	39	72.2	183	10	AV415207	AV415207 AV415207
36	39	72.2	201	14	D62821	D62821 HUM329C09B
c 37	39	72.2	231	9	AV036702	AV036702 AV036702
c 38	39	72.2	270	9	AV037125	AV037125 AV037125
c 39	39	72.2	274	9	AV036809	AV036809 AV036809
c 40	39	72.2	279	9	AV035994	AV035994 AV035994
c 41	39	72.2	281	9	AV037570	AV037570 AV037570
c 42	39	72.2	284	9	AV035331	AV035331 AV035331
c 43	39	72.2	284	10	BE523442	BE523442 M37A4STM
c 44	39	72.2	298	9	AV034428	AV034428 AV034428
c 45	39	72.2	306	9	AV225900	AV225900 AV225900

ALIGNMENTS

RESULT 1
BG216020/c
LOCUS BG216020 788 bp mRNA linear EST 21-APR-2001
DEFINITION RST35834 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG216020
VERSION BG216020.1 GI:13742169
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 788)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.
, Lerner,L., Costanzo,D., McElligott,K., Boozer,S., Mays,R., Smith

```
            ,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher
            ,J., Danzig,J. and Ducar,M.
  TITLE     Creation of genome-wide protein expression libraries using random
            activation of gene expression
  JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
  MEDLINE   21227151
COMMENT     Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 429.
FEATURES             Location/Qualifiers
     source          1. .788
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Athersys RAGE Library"
                     /cell_line="HT1080"
                     /note="See 'Creation of Genome-wide Protein Expression
                     Libraries using Random Activation of Gene Expression',
                     Nature Biotechnology, in press. Note that even though the
                     cell type indicated is HT1080, since a random activation
                     method was used, these sequence tags are not necessarily
                     expressed in HT1080 under normal circumstances."
BASE COUNT      193 a    216 c    198 g    181 t
ORIGIN

Alignment Scores:
Pred. No.:              124          Length:        788
Score:                  44.00        Matches:       8
Percent Similarity:     90.91%       Conservative:  2
Best Local Similarity:  72.73%       Mismatches:    1
Query Match:            81.48%       Indels:        0
DB:                     12           Gaps:          0

US-09-823-649A-3 (1-11) x BG216020 (1-788)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          |||||||||   |||:::::::|||||||||||||
Db    546 CTGTCCCAAAACCTTTCATTGCCTTACGAAGAA 514

RESULT 2
BI921437/c
LOCUS       BI921437                319 bp    mRNA    linear   EST 17-OCT-2001
DEFINITION  EST541340 tomato callus Lycopersicon esculentum cDNA clone
            cLEC73D21 5' end, mRNA sequence.
ACCESSION   BI921437
VERSION     BI921437.1  GI:16217465
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 319)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai
            ,J., Utterback,T., Van Aken,S., Ronning,C.M., Fraser,C.M., Martin
            ,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato callus tissue (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .319
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEC73D21"
                     /clone_lib="tomato callus"
                     /tissue_type="callus"
                     /dev_stage="25-40 days old"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Boyce Thompson Institute; sequencing: The
                     Institute for Genomic Research; cLEC - Cotyledons of
                     seedlings 7-10 days post-germination were excised, cut at
                     both ends and placed on MS medium with no selection. Mixed
                     callus was harvested at 25 and 40 days and included
                     undifferentiated masses. Tomato Callus EST Library"
BASE COUNT      108 a     49 c     89 g     73 t
ORIGIN

Alignment Scores:
Pred. No.:              90.3         Length:        319
Score:                  42.00        Matches:       8
Percent Similarity:     90.91%       Conservative:  2
Best Local Similarity:  72.73%       Mismatches:    1
Query Match:            77.78%       Indels:        0
DB:                     13           Gaps:          0

US-09-823-649A-3 (1-11) x BI921437 (1-319)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          |||||||||:::   |||||||||||||:::|||
Db     88 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 56

RESULT 3
BF050176/c
LOCUS       BF050176                467 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST435334 tomato developing/immature green fruit Lycopersicon
            esculentum cDNA clone cLEM16D8 5' sequence, mRNA sequence.
ACCESSION   BF050176
VERSION     BF050176.1  GI:10804072
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 467)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
            Liang,F., Hansen,I.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
            Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
            ,S.D.
  TITLE     Generation of ESTs from tomato fruit tissue, immature green
  JOURNAL   Unpublished (2000)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
FEATURES             Location/Qualifiers
     source          1. .467
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEM16D8"
                     /clone_lib="tomato developing/immature green fruit"
                     /tissue_type="fruit"
                     /dev_stage="immature green (5-35 days post-anthesis)"
                     /lab_host="SOLR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoRI;
                     Site_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
                     harvested at 7 day intervals through 35 dpa.  Equal masses
                     of tissue from each stage were combined (including seeds
                     and locules) prior to mRNA isolation."
BASE COUNT      149 a     72 c    139 g    107 t
```

```
ORIGIN

Alignment Scores:
Pred. No.:                153           Length:        467
Score:                    42.00         Matches:       8
Percent Similarity:       90.91%        Conservative:  2
Best Local Similarity:    72.73%        Mismatches:    1
Query Match:              77.78%        Indels:        0
DB:                       12            Gaps:          0

US-09-823-649A-3 (1-11) x BF050176 (1-467)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||:::   |||||||||||||:::|||
Db    326 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 294

RESULT 4
AW932704/c
LOCUS       AW932704                526 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST358547 tomato fruit mature green, TAMU Lycopersicon esculentum
            cDNA clone cLEF49J21 5', mRNA sequence.
ACCESSION   AW932704
VERSION     AW932704.1  GI:8108105
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 526)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
            Upton,J., Hansen?T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.
            , Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .526
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEF49J21"
                     /clone_lib="tomato fruit mature green, TAMU"
                     /tissue_type="fruit pericarp"
                     /dev_stage="mature green (3-5 days pre-ripening)"
                     /lab_host="SOLR"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; cLEF - Fruit were tagged at the 1cm stage and
                     harvested 3-5 days prior to  ripening. Fruit were cut in
                     half to verify the seeds were indeed  'immature' and the
                     seeds and locules were discarded prior to freezing the
                     pericarp"
BASE COUNT      169 a     91 c    139 g    127 t
ORIGIN

Alignment Scores:
Pred. No.:                180           Length:        526
Score:                    42.00         Matches:       8
Percent Similarity:       90.91%        Conservative:  2
Best Local Similarity:    72.73%        Mismatches:    1
Query Match:              77.78%        Indels:        0
DB:                       10            Gaps:          0

US-09-823-649A-3 (1-11) x AW932704 (1-526)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||:::   |||||||||||||:::|||
Db    177 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 145

RESULT 5
BI932551/c
LOCUS       BI932551                528 bp    mRNA    linear   EST 18-OCT-2001
DEFINITION  EST552440 tomato flower, 8 mm to preanthesis buds Lycopersicon
            esculentum cDNA clone cTOC23B21 5' end, mRNA sequence.
ACCESSION   BI932551
VERSION     BI932551.1  GI:16247023
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, buds 8 mm -
            preanthesis
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOC23B21"
                     /clone_lib="tomato flower, 8 mm to preanthesis buds"
                     /tissue_type="flower"
                     /dev_stage="buds 8mm to preanthesis"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research; Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT      168 a     91 c    146 g    123 t
ORIGIN

Alignment Scores:
Pred. No.:                181           Length:        528
Score:                    42.00         Matches:       8
Percent Similarity:       90.91%        Conservative:  2
Best Local Similarity:    72.73%        Mismatches:    1
Query Match:              77.78%        Indels:        0
DB:                       13            Gaps:          0

US-09-823-649A-3 (1-11) x BI932551 (1-528)

Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||:::   |||||||||||||:::|||
Db    240 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 208

RESULT 6
AW031837/c
LOCUS       AW031837                529 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST275291 tomato callus, TAMU Lycopersicon esculentum cDNA clone
            cLEC39F7 similar to ATP-dependent CLP protease ATP-binding subunit
            CLPA homolog, mRNA sequence.
ACCESSION   AW031837
VERSION     AW031837.1  GI:5890593
KEYWORDS    EST.
SOURCE      tomato.
```

```
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
            ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato callus tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .529
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEC39F7"
                     /clone_lib="tomato callus, TAMU"
                     /tissue_type="callus"
                     /dev_stage="25-40 days old"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni laboratory; cLEC - Cotyledons
                     of seedlings 7-10 days post-germination were excised, cut
                     at both ends and placed on MS medium with no selection.
                     Mixed callus was harvested at 25 and 40 days and included
                     undifferentiated masses. Tomato Callus EST Library"
BASE COUNT      169 a     93 c    139 g    128 t
ORIGIN

Alignment Scores:
Pred. No.:                181           Length:          529
Score:                    42.00         Matches:         8
Percent Similarity:       90.91%        Conservative:    2
Best Local Similarity:    72.73%        Mismatches:      1
Query Match:              77.78%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-3 (1-11) x AW031837 (1-529)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||:::   |||||||||||:::|||
Db     156 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 124

RESULT 7
AW399580/c
LOCUS       AW399580                537 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST310080 L. pennellii trichome, Cornell University Lycopersicon
            pennellii cDNA clone cLPT8E23 5', mRNA sequence.
ACCESSION   AW399580
VERSION     AW399580.1  GI:6918050
KEYWORDS    EST.
SOURCE      Lycopersicon pennellii.
  ORGANISM  Lycopersicon pennellii
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Lakey,J., Holt,I.E.,
            Liang,F., Hansen,T.S., Upton,J., Ronning,C.M., Craven,M.B., Fujii
            ,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin
            ,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from wild tomato (Lycopersicon pennellii)
            trichomes
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Lycopersicon pennellii"
                     /db_xref="taxon:28526"
                     /clone="cLPT8E23"
                     /clone_lib="L. pennellii trichome, Cornell University"
                     /tissue_type="trichome"
                     /dev_stage="mixed stages"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK-; Site_1: EcoR1; Site_2:
                     Xho1; Leaves of various stages were shaken in liquid
                     nitrogen, shearing off trichomes. This procedure yielded a
                     mixture of cells highly enriched for trichomes, with minor
                     contamination by other types of leaf cells."
BASE COUNT      163 a    104 c    139 g    131 t
ORIGIN

Alignment Scores:
Pred. No.:                185           Length:          537
Score:                    42.00         Matches:         8
Percent Similarity:       90.91%        Conservative:    2
Best Local Similarity:    72.73%        Mismatches:      1
Query Match:              77.78%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-3 (1-11) x AW399580 (1-537)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||:::   |||||||||||:::|||
Db      49 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 17

RESULT 8
AW930945/c
LOCUS       AW930945                543 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST356788 tomato fruit mature green, TAMU Lycopersicon esculentum
            cDNA clone cLEF42021 5', mRNA sequence.
ACCESSION   AW930945
VERSION     AW930945.1  GI:8106346
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
            Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.
            , Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .543
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEF42021"
                     /clone_lib="tomato fruit mature green, TAMU"
                     /tissue_type="fruit pericarp"
                     /dev_stage="mature green (3-5 days pre-ripening)"
                     /lab_host="SOLR"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
```

Xho1; cLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"

BASE COUNT 170 a 100 c 138 g 135 t
ORIGIN

Alignment Scores:

Pred. No.:	188	Length:	543
Score:	42.00	Matches:	8
Percent Similarity:	90.91%	Conservative:	2
Best Local Similarity:	72.73%	Mismatches:	1
Query Match:	77.78%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-3 (1-11) x AW930945 (1-543)

```
Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||:::   |||||||||||||:::|||
Db     87 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 55
```

RESULT 9
AW931411/c
LOCUS AW931411 558 bp mRNA linear EST 18-MAY-2001
DEFINITION EST357254 tomato fruit mature green, TAMU Lycopersicon esculentum cDNA clone cLEF45E19 5', mRNA sequence.
ACCESSION AW931411
VERSION AW931411.1 GI:8106812
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 558)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES Location/Qualifiers
source 1. .558
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF45E19"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2: Xho1; cLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"

BASE COUNT 176 a 102 c 145 g 135 t
ORIGIN

Alignment Scores:

Pred. No.:	195	Length:	558
Score:	42.00	Matches:	8
Percent Similarity:	90.91%	Conservative:	2
Best Local Similarity:	72.73%	Mismatches:	1
Query Match:	77.78%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-3 (1-11) x AW931411 (1-558)

```
Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||:::   |||||||||||||:::|||
Db    153 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 121
```

RESULT 10
AW931055/c
LOCUS AW931055 590 bp mRNA linear EST 18-MAY-2001
DEFINITION EST356898 tomato fruit mature green, TAMU Lycopersicon esculentum cDNA clone cLEF42H12 5', mRNA sequence.
ACCESSION AW931055
VERSION AW931055.1 GI:8106456
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 590)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES Location/Qualifiers
source 1. .590
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF42H12"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2: Xho1; cLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"

BASE COUNT 170 a 99 c 182 g 139 t
ORIGIN

Alignment Scores:

Pred. No.:	211	Length:	590
Score:	42.00	Matches:	8
Percent Similarity:	90.91%	Conservative:	2
Best Local Similarity:	72.73%	Mismatches:	1
Query Match:	77.78%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-3 (1-11) x AW931055 (1-590)

```
Qy      1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
          ||||||||||:::   |||||||||||||:::|||
Db    450 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 418
```

RESULT 11
AW737720/c
LOCUS AW737720 622 bp mRNA linear EST 18-MAY-2001
DEFINITION EST339147 tomato flower buds, anthesis, Cornell University Lycopersicon esculentum cDNA clone cTOD4G9 5', mRNA sequence.
ACCESSION AW737720

```
VERSION     AW737720.1  GI:7646665
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 622)
  AUTHORS   van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang
            ,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman
            ,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, anthesis
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .622
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOD4G9"
                     /clone_lib="tomato flower buds, anthesis, Cornell
                     University"
                     /tissue_type="flower"
                     /dev_stage="anthesis"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Tanksley;  Flower buds and flowers were
                     taken from greenhouse plants (4-8 wks old, TA496). They
                     were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT      191 a    116 c    158 g    157 t
ORIGIN

Alignment Scores:
Pred. No.:                227           Length:        622
Score:                    42.00         Matches:       8
Percent Similarity:       90.91%        Conservative:  2
Best Local Similarity:    72.73%        Mismatches:    1
Query Match:              77.78%        Indels:        0
DB:                       10            Gaps:          0

US-09-823-649A-3 (1-11) x AW737720 (1-622)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||:::   |||||||||||||:::|||
Db     154 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 122

RESULT 12
AW036109
LOCUS       AW036109                653 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST274485 tomato seed, TAMU Lycopersicon esculentum cDNA clone
            cLEE1M9 similar to ATP-dependent protease (CD4B), mRNA sequence.
ACCESSION   AW036109
VERSION     AW036109.1  GI:5894788
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 653)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato seed tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            3 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .653
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEE1M9"
                     /clone_lib="tomato seed, TAMU"
                     /tissue_type="seeds"
                     /dev_stage="quiescent seed"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLEE - Tomato Seed EST Library. Directionally cloned
                     cDNAs inserted into pBlueScript SK(-) at 5' end with
                     EcoR1 and 3' end with Xho1 site."
BASE COUNT      227 a    137 c    105 g    184 t
ORIGIN

Alignment Scores:
Pred. No.:                243           Length:        653
Score:                    42.00         Matches:       8
Percent Similarity:       90.91%        Conservative:  2
Best Local Similarity:    72.73%        Mismatches:    1
Query Match:              77.78%        Indels:        0
DB:                       10            Gaps:          0

US-09-823-649A-3 (1-11) x AW036109 (1-653)

Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||:::   |||||||||||||:::|||
Db     160 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 192

RESULT 13
BI933283/c
LOCUS       BI933283                743 bp    mRNA    linear   EST 18-OCT-2001
DEFINITION  EST553172 tomato flower, anthesis Lycopersicon esculentum cDNA
            clone cTOD16E14 5' end, mRNA sequence.
ACCESSION   BI933283
VERSION     BI933283.1  GI:16247755
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 743)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, anthesis (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .743
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOD16E14"
                     /clone_lib="tomato flower, anthesis"
                     /tissue_type="flower"
```

/dev_stage="anthesis"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Cornell University; sequencing: The Institute for Genomic Research; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."
BASE COUNT 216 a 135 c 218 g 174 t
ORIGIN

Alignment Scores:

Pred. No.:	290	Length:	743
Score:	42.00	Matches:	8
Percent Similarity:	90.91%	Conservative:	2
Best Local Similarity:	72.73%	Mismatches:	1
Query Match:	77.78%	Indels:	0
DB:	13	Gaps:	0

US-09-823-649A-3 (1-11) x BI933283 (1-743)

```
Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||:::   ||||||||||||:::|||
Db     509 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 477
```

RESULT 14
BG124823/c
LOCUS BG124823 759 bp mRNA linear EST 31-JAN-2001
DEFINITION EST470469 tomato shoot/meristem Lycopersicon esculentum cDNA clone cTOF6H13 5' sequence, mRNA sequence.
ACCESSION BG124823
VERSION BG124823.1 GI:12625011
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 759)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T., Hansen,C., Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES Location/Qualifiers
source 1. .759
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOF6H13"
/clone_lib="tomato shoot/meristem"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoR1; Site_2: Xho1; Small expanding leaves from the growing tip were taken from greenhouse plants (4-6wks old TA496). Tissue was immediately frozen in liquid nitrogen."
BASE COUNT 250 a 140 c 198 g 171 t
ORIGIN

Alignment Scores:

Pred. No.:	299	Length:	759
Score:	42.00	Matches:	8
Percent Similarity:	90.91%	Conservative:	2
Best Local Similarity:	72.73%	Mismatches:	1
Query Match:	77.78%	Indels:	0
DB:	12	Gaps:	0

US-09-823-649A-3 (1-11) x BG124823 (1-759)

```
Qy       1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||:::   ||||||||||||:::|||
Db      64 TTGTCCCAGCAACAAGCAATCCCATATCAAGAG 32
```

RESULT 15
AV043104
LOCUS AV043104 198 bp mRNA linear EST 23-NOV-1999
DEFINITION AV043104 Mus musculus adult C57BL/6J testis Mus musculus cDNA clone 1700026C12, mRNA sequence.
ACCESSION AV043104
VERSION AV043104.2 GI:4862769
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 198)
AUTHORS Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K., Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T., Hara,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M., Kawai,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H., Owa,C., Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tateno,M., Tomaru,Y., Tominaga,N., Watanabe,S., Yagame,M., Yamamura,T., Yokota,T., Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs
JOURNAL Unpublished (1999)
COMMENT Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .198
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1700026C12"
/clone_lib="Mus musculus adult C57BL/6J testis"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TCTTACCAATCTGAAGTGGGAGCCGCCCGCAATGGTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 61 a 48 c 33 g 56 t
ORIGIN

Alignment Scores:

Pred. No.:	74.3	Length:	198
Score:	41.00	Matches:	8
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	1
Query Match:	75.93%	Indels:	0

DB: 9 Gaps: 0

US-09-823-649A-3 (1-11) x AV043104 (1-198)

```
Qy       3 GlnGluLeuAlaIleProTyrGluGlu 11
           ||||||||||||   ||||||||||||
Db      17 CAAGAACTCGCCTATCCTTATGAAGAG 43
```

Search completed: January 21, 2003, 12:24:16
Job time : 1171.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:02:09 ; Search time 28.1429 Seconds
(without alignments)
52.083 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSXELSIPYEE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48	96.0	11	23	AAB47794	Native DNA polymer
2	48	96.0	11	23	AAM48260	Native DNA polymer
3	48	96.0	11	23	AAM48263	Native DNA polymer
4	48	96.0	11	23	AAM48265	Native DNA polymer
5	48	96.0	545	13	AAR23157	Mutant thermostabl
6	48	96.0	632	13	AAR23156	Mutant thermostabl
7	48	96.0	680	13	AAR23155	Mutant thermostabl
8	48	96.0	758	13	AAR23154	Mutant thermostabl
9	48	96.0	788	13	AAR23153	Mutant thermostabl
10	48	96.0	830	13	AAR23152	Mutant thermostabl
11	48	96.0	830	16	AAR76060	Tfil DNA-polymeras
12	48	96.0	831	16	AAR64273	T. flavus DNA-poly
13	48	96.0	831	19	AAW79961	Thermus flavus DNA
14	48	96.0	831	20	AAW80428	DNA polymerase enz
15	48	96.0	845	16	AAR72875	Tsps17 polymerase.
16	45	90.0	11	23	AAB47793	Native DNA polymer
17	45	90.0	11	23	AAM48259	Native DNA polymer
18	45	90.0	11	23	AAM48261	Native DNA polymer
19	45	90.0	11	23	AAM48262	Native DNA polymer
20	45	90.0	11	23	AAM48264	Native DNA polymer
21	45	90.0	528	17	AAR96267	Mutant Thermus aqu
22	45	90.0	544	13	AAR23145	Mutant thermostabl
23	45	90.0	544	13	AAR23163	Mutant thermostabl
24	45	90.0	544	13	AAR23168	Mutant thermostabl
25	45	90.0	552	22	AAE09304	Thermus aquaticus
26	45	90.0	553	21	AAY44352	Thermostable DNA d
27	45	90.0	554	16	AAR66209	Novel thermostable
28	45	90.0	559	21	AAY44353	Thermostable DNA d
29	45	90.0	560	17	AAR99544	Thermus flavus DNA
30	45	90.0	560	17	AAR96205	DNA polymerase I e
31	45	90.0	561	18	AAW09315	Mutant Taq polymer
32	45	90.0	561	18	AAW09317	Mutant Taq polymer
33	45	90.0	561	20	AAY00886	FY2 polymerase. T
34	45	90.0	561	20	AAY00888	FY4 polymerase. T
35	45	90.0	562	18	AAW09316	Mutant Taq polymer
36	45	90.0	562	20	AAY00887	FY3 polymerase. T
37	45	90.0	562	22	AAE09305	Thermus thermophil
38	45	90.0	562	22	AAE09310	Thermus aquaticus
39	45	90.0	597	17	AAR99543	Thermus flavus DNA
40	45	90.0	597	17	AAR96204	DNA polymerase I e
41	45	90.0	604	22	AAU00574	Chimeric CauTaq DN
42	45	90.0	604	22	AAU00575	Chimeric CauTaq DN
43	45	90.0	631	13	AAR23144	Mutant thermostabl
44	45	90.0	632	13	AAR23162	Mutant thermostabl
45	45	90.0	632	13	AAR23167	Mutant thermostabl

ALIGNMENTS

RESULT 1
AAB47794
ID AAB47794 standard; peptide; 11 AA.
XX
AC AAB47794;
XX
DT 25-MAR-2002 (first entry)
XX
DE Native DNA polymerase motif #4.
XX
KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KW transverse transcription reaction; fluorescein; cyanine; thermoactive;
KW dye; amplification.
XX
OS Thermus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 3
FT /label= Gln, Gly
XX
PN EP1152062-A2.
XX
PD 07-NOV-2001.
XX
PF 12-APR-2001; 2001EP-0109341.
XX
PR 18-APR-2000; 2000US-198336P.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schoenbrunner NJ, Wang AM;
XX

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds
(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649A-2
Perfect score: 47
Sequence: 1 LSXELXIPYEE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24626/app_query.fasta_1.1393
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649_@CGN_1_1_77_@runat_21012003_093151_24626
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	44	93.6	1600	9	US-10-033-297-12	Sequence 12, Appl
2	44	93.6	1600	9	US-10-081-806-12	Sequence 12, Appl
3	44	93.6	2496	9	US-10-033-297-2	Sequence 2, Appli
4	44	93.6	2496	9	US-10-081-806-2	Sequence 2, Appli
5	44	93.6	2502	9	US-10-033-297-7	Sequence 7, Appli
6	44	93.6	2502	9	US-10-033-297-21	Sequence 21, Appl
7	44	93.6	2502	9	US-10-081-806-7	Sequence 7, Appli
8	44	93.6	2502	9	US-10-081-806-21	Sequence 21, Appl
9	44	93.6	2504	9	US-10-033-297-3	Sequence 3, Appli
10	44	93.6	2504	9	US-10-081-806-3	Sequence 3, Appli
11	44	93.6	2505	9	US-10-033-297-65	Sequence 65, Appl
12	44	93.6	2505	9	US-10-033-297-68	Sequence 68, Appl
13	44	93.6	2505	9	US-10-033-297-70	Sequence 70, Appl
14	44	93.6	2506	9	US-10-033-297-1	Sequence 1, Appli
15	44	93.6	2506	9	US-10-081-806-1	Sequence 1, Appli
16	44	93.6	2511	10	US-09-777-430A-9	Sequence 9, Appli
17	44	93.6	2511	10	US-09-777-430A-12	Sequence 12, Appl
18	44	93.6	2526	10	US-09-777-430A-14	Sequence 14, Appl
19	44	93.6	2526	10	US-09-777-430A-19	Sequence 19, Appl
20	44	93.6	2526	10	US-09-777-430A-22	Sequence 22, Appl
21	44	93.6	2526	10	US-09-777-430A-25	Sequence 25, Appl
22	44	93.6	2626	9	US-09-972-834-1	Sequence 1, Appli
23	33	70.2	378	10	US-09-734-017A-57	Sequence 57, Appl
24	33	70.2	402	10	US-09-974-300-3201	Sequence 3201, Ap
25	33	70.2	1497	10	US-09-735-169A-3	Sequence 3, Appli
26	33	70.2	1497	10	US-09-735-171A-3	Sequence 3, Appli
27	33	70.2	3452	10	US-09-735-169A-1	Sequence 1, Appli
28	33	70.2	3452	10	US-09-735-171A-1	Sequence 1, Appli
C 29	33	70.2	32768	10	US-09-070-927A-128	Sequence 128, App
C 30	32	68.1	69	10	US-09-850-492-2	Sequence 2, Appli
31	32	68.1	275	10	US-09-878-574-5363	Sequence 5363, Ap
C 32	32	68.1	286	10	US 09-923-876-6139	Sequence 6139, Ap
33	32	68.1	342	10	US-09-983-965-142	Sequence 142, App
34	32	68.1	467	10	US-09-920-300A-911	Sequence 911, App
35	32	68.1	467	12	US-10-033-528-911	Sequence 911, App
C 36	32	68.1	1035	9	US-10-002-344A-85	Sequence 85, Appl
37	32	68.1	1524	10	US-09-962-436-42	Sequence 42, Appl
38	32	68.1	1620	9	US-09-938-842A-960	Sequence 960, App
39	32	68.1	1807	10	US-09-880-107-3712	Sequence 3712, Ap
40	32	68.1	2088	9	US-10-033-297-10	Sequence 10, Appl
41	32	68.1	2088	9	US-10-081-806-10	Sequence 10, Appl
42	32	68.1	2603	9	US-09-764-868-80	Sequence 80, Appl
43	32	68.1	3203	10	US-09-925-302-142	Sequence 142, App
44	32	68.1	3546	10	US-09-921-232-14	Sequence 14, Appl
45	32	68.1	3546	10	US-09-921-232-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-10-033-297-12
; Sequence 12, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; Lyamichev, Victor I
; Mast, Andrea L.
; Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/033,297
; FILING DATE: 12-No. US20020187486A1-2001
; CLASSIFICATION: <Unknown>

```
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;    INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1600 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-033-297-12

Alignment Scores:
Pred. No.:                0.559          Length:          1600
Score:                    44.00          Matches:         9
Percent Similarity:       81.82%         Conservative:    0
Best Local Similarity:    81.82%         Mismatches:      2
Query Match:              93.62%         Indels:          0
DB:                       9              Gaps:            0

US-09-823-649A-2 (1-11) x US-10-033-297-12 (1-1600)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   |||||||||||||||
Db    1133 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 1165

RESULT 2
US-10-081-806-12
; Sequence 12, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;        APPLICANT: Prudent, James R.
;                   Hall, Jeff G.
;                   Lyamichev, Victor I.
;        TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;        NUMBER OF SEQUENCES: 69
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;    INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1600 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-081-806-12

Alignment Scores:
Pred. No.:                0.559          Length:          1600
Score:                    44.00          Matches:         9
Percent Similarity:       81.82%         Conservative:    0
Best Local Similarity:    81.82%         Mismatches:      2
Query Match:              93.62%         Indels:          0
DB:                       9              Gaps:            0

US-09-823-649A-2 (1-11) x US-10-081-806-12 (1-1600)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   |||||||||||||||
Db    1133 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 1165

RESULT 3
US-10-033-297-2
; Sequence 2, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
```

```
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2496 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-033-297-2

Alignment Scores:
Pred. No.:                0.908          Length:          2496
Score:                    44.00          Matches:         9
Percent Similarity:       81.82%         Conservative:    0
Best Local Similarity:    81.82%         Mismatches:      2
Query Match:              93.62%         Indels:          0
DB:                       9              Gaps:            0

US-09-823-649A-2 (1-11) x US-10-033-297-2 (1-2496)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   ||||||||||||||||
Db    2029 CTCTCCGGGGAGCTTTCCATCCCCTACGAGGAG 2061

RESULT 4
US-10-081-806-2
; Sequence 2, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;       APPLICANT: Prudent, James R.
;                  Hall, Jeff G.
;                  Lyamichev, Victor I.
;       TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;       NUMBER OF SEQUENCES: 69
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 2496 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-081-806-2

Alignment Scores:
Pred. No.:                0.908          Length:          2496
Score:                    44.00          Matches:         9
Percent Similarity:       81.82%         Conservative:    0
Best Local Similarity:    81.82%         Mismatches:      2
Query Match:              93.62%         Indels:          0
DB:                       9              Gaps:            0

US-09-823-649A-2 (1-11) x US-10-081-806-2 (1-2496)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   ||||||||||||||||
Db    2029 CTCTCCGGGGAGCTTTCCATCCCCTACGAGGAG 2061

RESULT 5
US-10-033-297-7
; Sequence 7, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCI/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
```

```
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-033-297-7

Alignment Scores:
Pred. No.:              0.91          Length:          2502
Score:                  44.00         Matches:         9
Percent Similarity:     81.82%        Conservative:    0
Best Local Similarity:  81.82%        Mismatches:      2
Query Match:            93.62%        Indels:          0
DB:                     9             Gaps:            0

US-09-823-649A-2 (1-11) x US-10-033-297-7 (1-2502)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||  ||||||||||||||||
Db   2035 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2067

RESULT 6
US-10-033-297-21
; Sequence 21, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 21:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-033-297-21

Alignment Scores:
Pred. No.:              0.91          Length:          2502
Score:                  44.00         Matches:         9
Percent Similarity:     81.82%        Conservative:    0
Best Local Similarity:  81.82%        Mismatches:      2
Query Match:            93.62%        Indels:          0
DB:                     9             Gaps:            0

US-09-823-649A-2 (1-11) x US-10-033-297-21 (1-2502)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||  ||||||||||||||||
Db   2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 7
US-10-081-806-7
; Sequence 7, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;        APPLICANT: Prudent, James R.
;                   Hall, Jeff G.
;                   Lyamichev, Victor I.
;        TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;        NUMBER OF SEQUENCES: 69
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
```

```
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-081-806-7

Alignment Scores:
Pred. No.:                0.91          Length:          2502
Score:                    44.00         Matches:         9
Percent Similarity:       81.82%        Conservative:    0
Best Local Similarity:    81.82%        Mismatches:      2
Query Match:              93.62%        Indels:          0
DB:                       9             Gaps:            0

US-09-823-649A-2 (1-11) x US-10-081-806-7 (1-2502)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   ||||||||||||||||
Db    2035 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2067

RESULT 8
US-10-081-806-21
; Sequence 21, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;        APPLICANT: Prudent, James R.
;                   Hall, Jeff G.
;                   Lyamichev, Victor I.
;        TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;        NUMBER OF SEQUENCES: 69
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 21:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2502 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-081-806-21

Alignment Scores:
Pred. No.:                0.91          Length:          2502
Score:                    44.00         Matches:         9
Percent Similarity:       81.82%        Conservative:    0
Best Local Similarity:    81.82%        Mismatches:      2
Query Match:              93.62%        Indels:          0
DB:                       9             Gaps:            0

US-09-823-649A-2 (1-11) x US-10-081-806-21 (1-2502)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   ||||||||||||||||
Db    2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 9
US-10-033-297-3
; Sequence 3, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,038
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
```

```
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2504 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-033-297-3

Alignment Scores:
Pred. No.:                0.911          Length:          2504
Score:                    44.00          Matches:         9
Percent Similarity:       81.82%         Conservative:    0
Best Local Similarity:    81.82%         Mismatches:      2
Query Match:              93.62%         Indels:          0
DB:                       9              Gaps:            0

US-09-823-649A-2 (1-11) x US-10-033-297-3 (1-2504)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   |||||||||||||||
Db    2038 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2070

RESULT 10
US-10-081-806-3
; Sequence 3, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;        APPLICANT: Prudent, James R.
;                   Hall, Jeff G.
;                   Lyamichev, Victor I.
;        TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;        NUMBER OF SEQUENCES: 69
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Patentin Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/081,806
;             FILING DATE: 22-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/756,386
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2504 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-081-806-3

Alignment Scores:
Pred. No.:                0.911          Length:          2504
Score:                    44.00          Matches:         9
Percent Similarity:       81.82%         Conservative:    0
Best Local Similarity:    81.82%         Mismatches:      2
Query Match:              93.62%         Indels:          0
DB:                       9              Gaps:            0

US-09-823-649A-2 (1-11) x US-10-081-806-3 (1-2504)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   |||||||||||||||
Db    2038 CTCTCCCAGGAGCTTGCCATCCCCTACGAGGAG 2070

RESULT 11
US-10-033-297-65
; Sequence 65, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;        APPLICANT: Hall, Jeff G.
;                   Lyamichev, Victor I.
;                   Mast, Andrea L.
;                   Brow, Mary Ann D.
;        TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                            Sequential Invasive Cleavages
;        NUMBER OF SEQUENCES: 163
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Medlen & Carroll, LLP
;             STREET: 220 Montgomery Street, Suite 2200
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: United States Of America
;             ZIP: 94104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Patentin Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/033,297
;             FILING DATE: 12-No. US20020187486A1-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,597
;             FILING DATE: 09-Jul-1999
;             APPLICATION NUMBER: US/08/823,516
;             FILING DATE: 24-MAR-1997
;             APPLICATION NUMBER: PCT/US97/01072
;             FILING DATE: 21-JAN-1997
;             APPLICATION NUMBER: US 08/759,648
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/758,314
;             FILING DATE: 02-DEC-1996
;             APPLICATION NUMBER: US 08/756,386
;             FILING DATE: 29-NOV-1996
;             APPLICATION NUMBER: US 08/682,853
;             FILING DATE: 12-JUL-1996
;             APPLICATION NUMBER: US 08/599,491
;             FILING DATE: 24-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ingolia, Diane E.
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02736
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 65:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2505 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
```

```
;       MOLECULE TYPE: DNA (genomic)
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..2499
;       SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-10-033-297-65

Alignment Scores:
Pred. No.:              0.911          Length:         2505
Score:                  44.00          Matches:        9
Percent Similarity:     81.82%         Conservative:   0
Best Local Similarity:  81.82%         Mismatches:     2
Query Match:            93.62%         Indels:         0
DB:                     9              Gaps:           0

US-09-823-649A-2 (1-11) x US-10-033-297-65 (1-2505)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   ||||||||||||||||
Db    2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 12
US-10-033-297-68
; Sequence 68, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/033,297
;            FILING DATE: 12-No. US20020187486A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
;            FILING DATE: 21-JAN-1997
;            APPLICATION NUMBER: US 08/759,038
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/758,314
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/756,386
;            FILING DATE: 29-NOV-1996
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 68:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2505 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: other nucleic acid
;            DESCRIPTION: /desc = "DNA"
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..2499
;       SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-10-033-297-68

Alignment Scores:
Pred. No.:              0.911          Length:         2505
Score:                  44.00          Matches:        9
Percent Similarity:     81.82%         Conservative:   0
Best Local Similarity:  81.82%         Mismatches:     2
Query Match:            93.62%         Indels:         0
DB:                     9              Gaps:           0

US-09-823-649A-2 (1-11) x US-10-033-297-68 (1-2505)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   ||||||||||||||||
Db    2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 13
US-10-033-297-70
; Sequence 70, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/033,297
;            FILING DATE: 12-No. US20020187486A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
;            FILING DATE: 21-JAN-1997
;            APPLICATION NUMBER: US 08/759,038
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/758,314
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/756,386
;            FILING DATE: 29-NOV-1996
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
```

```
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 70:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2505 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: other nucleic acid
;            DESCRIPTION: /desc = "DNA"
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..2499
;       SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-10-033-297-70

Alignment Scores:
Pred. No.:              0.911          Length:         2505
Score:                  44.00          Matches:        9
Percent Similarity:     81.82%         Conservative:   0
Best Local Similarity:  81.82%         Mismatches:     2
Query Match:            93.62%         Indels:         0
DB:                     9              Gaps:           0

US-09-823-649A-2 (1-11) x US-10-033-297-70 (1-2505)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||.. ||||||||||||||
Db   2035 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2067

RESULT 14
US-10-033-297-1
; Sequence 1, Application US/10033297
; Publication No. US20020187486A1
;   GENERAL INFORMATION:
;       APPLICANT: Hall, Jeff G.
;                  Lyamichev, Victor I.
;                  Mast, Andrea L.
;                  Brow, Mary Ann D.
;       TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                           Sequential Invasive Cleavages
;       NUMBER OF SEQUENCES: 163
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/033,297
;            FILING DATE: 12-No. US20020187486A1-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/350,597
;            FILING DATE: 09-Jul-1999
;            APPLICATION NUMBER: US/08/823,516
;            FILING DATE: 24-MAR-1997
;            APPLICATION NUMBER: PCT/US97/01072
;            FILING DATE: 21-JAN-1997
;            APPLICATION NUMBER: US 08/759,038
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/758,314
;            FILING DATE: 02-DEC-1996
;            APPLICATION NUMBER: US 08/756,386
;            FILING DATE: 29-NOV-1996
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
;            REGISTRATION NUMBER: 40,027
;            REFERENCE/DOCKET NUMBER: FORS-02736
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (415) 705-8410
;            TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2506 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-033-297-1

Alignment Scores:
Pred. No.:              0.912          Length:         2506
Score:                  44.00          Matches:        9
Percent Similarity:     81.82%         Conservative:   0
Best Local Similarity:  81.82%         Mismatches:     2
Query Match:            93.62%         Indels:         0
DB:                     9              Gaps:           0

US-09-823-649A-2 (1-11) x US-10-033-297-1 (1-2506)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   ||||||||||||||
Db   2032 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2064

RESULT 15
US-10-081-806-1
; Sequence 1, Application US/10081806
; Publication No. US20020197623A1
;   GENERAL INFORMATION:
;       APPLICANT: Prudent, James R.
;                  Hall, Jeff G.
;                  Lyamichev, Victor I.
;       TITLE OF INVENTION: Invasive Cleavge Of Nucleic Acids
;       NUMBER OF SEQUENCES: 69
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Medlen & Carroll, LLP
;            STREET: 220 Montgomery Street, Suite 2200
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: United States Of America
;            ZIP: 94104
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/081,806
;            FILING DATE: 22-Feb-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/756,386
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: US 08/682,853
;            FILING DATE: 12-JUL-1996
;            APPLICATION NUMBER: US 08/599,491
;            FILING DATE: 24-JAN-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ingolia, Diane E.
```

```
;             REGISTRATION NUMBER: 40,027
;             REFERENCE/DOCKET NUMBER: FORS-02564
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 705-8410
;             TELEFAX: (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2506 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-081-806-1

Alignment Scores:
Pred. No.:              0.912          Length:          2506
Score:                  44.00          Matches:         9
Percent Similarity:     81.82%         Conservative:    0
Best Local Similarity:  81.82%         Mismatches:      2
Query Match:            93.62%         Indels:          0
DB:                     9              Gaps:            0

US-09-823-649A-2 (1-11) x US-10-081-806-1 (1-2506)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   ||||||||||||||||
Db    2032 CTCTCCCAGGAGCTAGCCATCCCTTACGAGGAG 2064


Search completed: January 21, 2003, 10:07:41
Job time : 31.5714 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds
(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649A-2
Perfect score: 47
Sequence: 1 LSXELXIPYEE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24616/app_query.fasta_1.1393
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649_@CGN_1_1_2820_@runat_21012003_093151_24616 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		41	87.2	447	13	BI929077	BI929077 EST548966
2		41	87.2	480	12	BF050758	BF050758 EST435916
3		41	87.2	495	13	BI129400	BI129400 G090P24Y
4		41	87.2	507	9	AI895448	AI895448 EST264891
5		41	87.2	534	14	BQ489763	BQ489763 05-E9425-
6		41	87.2	592	10	BE434433	BE434433 EST405511
7	c	41	87.2	629	14	BQ990012	BQ990012 QGF19D10.
8		41	87.2	728	14	BQ856219	BQ856219 QGB28014.
9		41	87.2	768	13	BM411779	BM411779 EST586106
10		40	85.1	658	14	BQ404238	BQ404238 GA__Ed006
11		38	80.9	183	10	AV415207	AV415207 AV415207
12	c	38	80.9	249	9	AV035497	AV035497 AV035497
13		38	80.9	411	10	AV428667	AV428667 AV428667
14	c	38	80.9	416	12	BF758509	BF758509 MR4-CT053
15		38	80.9	426	10	AV415722	AV415722 AV415722
16	c	37	78.7	231	9	AV036702	AV036702 AV036702
17	c	37	78.7	270	9	AV037125	AV037125 AV037125
18	c	37	78.7	274	9	AV036809	AV036809 AV036809
19	c	37	78.7	279	9	AV035994	AV035994 AV035994
20	c	37	78.7	281	9	AV037570	AV037570 AV037570
21	c	37	78.7	284	9	AV035331	AV035331 AV035331
22	c	37	78.7	298	9	AV034428	AV034428 AV034428
23	c	37	78.7	306	9	AV225900	AV225900 AV225900
24	c	37	78.7	319	9	AV036154	AV036154 AV036154
25	c	37	78.7	345	9	AI152562	AI152562 ud92t06.r
26		37	78.7	516	13	BM249366	BM249366 K0832C03-
27		37	78.7	562	13	BM250725	BM250725 K0854C08-
28	c	37	78.7	682	14	BQ185049	BQ185049 UI-E-EJ1-
29	c	37	78.7	788	12	BG216020	BG216020 RST35834
30	c	36	76.6	274	9	AV223368	AV223368 AV223368
31	c	36	76.6	438	13	BM482565	BM482565 535491 MA
32	c	36	76.6	571	12	BG834832	BG834832 353186 MA
33		36	76.6	587	13	BJ315731	BJ315731 BJ315731
34	c	36	76.6	721	13	BJ321189	BJ321189 BJ321189
35		36	76.6	760	13	BM291026	BM291026 EST577560
36	c	35	74.5	274	9	AV223298	AV223298 AV223298
37	c	35	74.5	292	9	AV034586	AV034586 AV034586
38	c	35	74.5	353	13	BI680984	BI680984 460228 MA
39	c	35	74.5	440	9	AI850988	AI850988 UI-M-BH0-
40	c	35	74.5	482	9	AA701609	AA701609 zi41e09.s
41	c	35	74.5	493	14	R39333	R39333 yd01h05.s1
42		35	74.5	565	10	BE656341	BE656341 UI-M-BH0-
43	c	35	74.5	647	12	BF701996	BF701996 MI-P-A1-a
44		35	74.5	650	17	AZ826009	AZ826009 2M0101C05
45	c	35	74.5	664	14	BQ139058	BQ139058 NF010E06P

ALIGNMENTS

RESULT 1
BI929077
LOCUS BI929077 447 bp mRNA linear EST 18-OCT-2001
DEFINITION EST548966 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA clone cTOB27L17 5' end, mRNA sequence.
ACCESSION BI929077
VERSION BI929077.1 GI:16241891
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 447)

```
  AUTHORS   van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
            Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
            Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
  TITLE     Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOB27L17"
                     /clone_lib="tomato flower, 3 - 8 mm buds"
                     /tissue_type="flower"
                     /dev_stage="3-8mm buds"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research. Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT      118 a     92 c     89 g    148 t
ORIGIN

Alignment Scores:
Pred. No.:                96.4 *          Length:         447
Score:                    41.00           Matches:        7
Percent Similarity:       81.82%          Conservative:   2
Best Local Similarity:    63.64%          Mismatches:     2
Query Match:              87.23%          Indels:         0
DB:                       13              Gaps:           0

US-09-823-649A-2 (1-11) x BI929077 (1-447)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   :::||||||||||:::
Db    331 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGGAT 363

RESULT 2
BF050758
LOCUS       BF050758                480 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  UST435916 tomato developing/immature green fruit Lycopersicon
            esculentum cDNA clone cLEM19022 5' sequence, mRNA sequence.
ACCESSION   BF050758
VERSION     BF050758.1  GI:10804654
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
            Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
            Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
            ,S.D.
  TITLE     Generation of ESTs from tomato fruit tissue, immature green
  JOURNAL   Unpublished (2000)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEM19022"
                     /clone_lib="tomato developing/immature green fruit"
                     /tissue_type="fruit"
                     /dev_stage="immature green (5-35 days post-anthesis)"
                     /lab_host="SOLR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoRI;
                     Site_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
                     harvested at 7 day intervals through 35 dpa.  Equal masses
                     of tissue from each stage were combined (including seeds
                     and locules) prior to mRNA isolation."
BASE COUNT      116 a    102 c    119 g    143 t
ORIGIN

Alignment Scores:
Pred. No.:                106             Length:         480
Score:                    41.00           Matches:        7
Percent Similarity:       81.82%          Conservative:   2
Best Local Similarity:    63.64%          Mismatches:     2
Query Match:              87.23%          Indels:         0
DB:                       12              Gaps:           0

US-09-823-649A-2 (1-11) x BF050758 (1-480)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   :::||||||||||:::
Db    261 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGGAT 293

RESULT 3
BI129400
LOCUS       BI129400                495 bp    mRNA    linear   EST 31-DEC-2001
DEFINITION  G090P24Y Populus cambium cDNA library Populus tremula x Populus
            tremuloides cDNA, mRNA sequence.
ACCESSION   BI129400
VERSION     BI129400.1  GI:18013371
KEYWORDS    EST.
SOURCE      Populus tremula x Populus tremuloides.
  ORGANISM  Populus tremula x Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   Hertzberg,M., Aspeborg,H., Erlandsson,R., Bjorkbacka,H., Hiltonen
            ,T., Karlsson,J., Teeri,T., Gustafsson,P., Bahlerao,R., Jansson,S.,
            Nilsson,O., Sundberg,B., Nilsson,P., Uhlen,M., Sandberg,G. and
            Lundeberg,J.
  TITLE     Gene expression in Populus
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Erlandsson R
            Department of Biotechnology
            Royal Institute of Technology
            Teknikringen 30, Stockholm S-10044, Sweden
            Tel: 46 8 790 8287
            Fax: 46 8 245452
            Email: rikerl@biochem.kth.se.
FEATURES             Location/Qualifiers
     source          1. .495
                     /organism="Populus tremula x Populus tremuloides"
                     /db_xref="taxon:47664"
                     /clone_lib="Populus cambium cDNA library"
                     /note="Organ: cambium"
BASE COUNT      109 a    146 c    117 g    123 t
ORIGIN

Alignment Scores:
Pred. No.:                110             Length:         495
Score:                    41.00           Matches:        7
Percent Similarity:       81.82%          Conservative:   2
Best Local Similarity:    63.64%          Mismatches:     2
Query Match:              87.23%          Indels:         0
DB:                       13              Gaps:           0
```

```
US-09-823-649A-2 (1-11) x BI129400 (1-495)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   :::||||||||||:::
Db     59 CTATCTGCCGAGTTGTCCGTTCCTTATGAAGAT 91

RESULT 4
AI895448
LOCUS       AI895448                507 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST264891 tomato callus, TAMU Lycopersicon esculentum cDNA clone
            cLEC7H2, mRNA sequence.
ACCESSION   AI895448
VERSION     AI895448.1  GI:5601350
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 507)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning
            ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato callus tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequencer
FEATURES             Location/Qualifiers
     source          1. .507
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEC7H2"
                     /clone_lib="tomato callus, TAMU"
                     /tissue_type="callus"
                     /dev_stage="25-40 days old"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni laboratory; cLEC - Cotyledons
                     of seedlings 7-10 days post-germination were excised, cut
                     at both ends and placed on MS medium with no selection.
                     Mixed callus was harvested at 25 and 40 days and included
                     undifferentiated masses. Tomato Callus EST Library"
BASE COUNT      125 a    138 c    114 g    130 t
ORIGIN

Alignment Scores:
Pred. No.:                114           Length:         507
Score:                    41.00         Matches:        7
Percent Similarity:       81.82%        Conservative:   2
Best Local Similarity:    63.64%        Mismatches:     2
Query Match:              87.23%        Indels:         0
DB:                       9             Gaps:           0

US-09-823-649A-2 (1-11) x AI895448 (1-507)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   :::||||||||||:::
Db     66 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGGAT 98

RESULT 5
BQ489763
LOCUS       BQ489763                534 bp    mRNA    linear   EST 07-JUN-2002
DEFINITION  05-E9425-006-009-I01-T3 Sugar beet MPIZ-ADIS-006 Lambda Zap II
            library Beta vulgaris cDNA clone I-1-9, mRNA sequence.
ACCESSION   BQ489763
VERSION     BQ489763.1  GI:21334383
KEYWORDS    EST.
SOURCE      Beta vulgaris.
  ORGANISM  Beta vulgaris
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
REFERENCE   1  (bases 1 to 534)
  AUTHORS   Bellin,D., Werber,M., Theis,T., Weisshaar,B. and Schneider,K.
  TITLE     EST sequencing, annotation and macroarray expression analysis of
            more than 4000 sugar beet cDNAs identifies genes with root-specific
            expression pattern
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Weisshaar B
            ADIS DNA core facility at MPIZ
            Max-Planck-Institute for Plant Breeding Research
            Carl-von-Linne Weg 10, 50829 Koeln, Germany
            Fax: 00492215062851
            Email: weisshaa@mpiz-koeln.mpg.de
            Seq primer: T3 'AATTAACCCTCACTAAAGGG'
            High quality sequence stop: 534.
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Beta vulgaris"
                     /db_xref="taxon:161934"
                     /clone="I-1-9"
                     /clone_lib="Sugar beet MPIZ-ADIS-006 Lambda Zap II
                     library"
                     /dev_stage="4 week old pot-grown plants"
                     /note="Organ: shoot and root; Vector: pBluescript SK- from
                     lambda ZAP II; cDNA (lambda ZAP-II) library from sugar
                     beet, whole plant mRNA, Prepared using the Stratagene
                     UniZAP cDNA kit, cloning sites EcoRI-XhoI, primer sites
                     and orientation:
                     rev-T3-SacI-SK-EcoRI-GGCACGAGG-5pr-cDNA-polyA-XhoI-KpnI-T7
                     -uni"
BASE COUNT      143 a    112 c    127 g    152 t
ORIGIN

Alignment Scores:
Pred. No.:                122           Length:         534
Score:                    41.00         Matches:        7
Percent Similarity:       81.82%        Conservative:   2
Best Local Similarity:    63.64%        Mismatches:     2
Query Match:              87.23%        Indels:         0
DB:                       14            Gaps:           0

US-09-823-649A-2 (1-11) x BQ489763 (1-534)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   :::||||||||||:::
Db    281 CTCTCGGCTGAACTCTCAGTTCCGTATGAGGAT 317

RESULT 6
BE434433
LOCUS       BE434433                592 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST405511 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
            clone cLEG17C9, mRNA sequence.
ACCESSION   BE434433
VERSION     BE434433.1  GI:9432276
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
            Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
            Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
            ,S.D.
  TITLE     Generation of ESTs from tomato fruit tissue, breaker stage
  JOURNAL   Unpublished (2000)
```

```
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .592
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEG17C9"
                     /clone_lib="tomato breaker fruit, TIGR"
                     /tissue_type="Pericarp"
                     /dev_stage="breaker"
                     /lab_host="SOLR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoRI;
                     Site_2: XhoI; Fruit were harvested at the breaker stage
                     (first sign of lycopene accumulation on the blossom end of
                     the fruit).  Fruit were cut in half and the seeds and
                     locules were discarded prior to freezing the pericarp."
BASE COUNT      140 a    149 c    137 g    166 t
ORIGIN

Alignment Scores:
Pred. No.:              139          Length:         592
Score:                  41.00        Matches:        7
Percent Similarity:     81.82%       Conservative:   2
Best Local Similarity:  63.64%       Mismatches:     2
Query Match:            87.23%       Indels:         0
DB:                     10           Gaps:           0

US-09-823-649A-2 (1-11) x BE434433 (1-592)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   :::||||||||||:::
Db    194 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGGAT 226

RESULT 7
BQ990012/c
LOCUS       BQ990012                629 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  QGF19D10.yg.ab1 QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
            QGF19D10, mRNA sequence.
ACCESSION   BQ990012
VERSION     BQ990012.1  GI:22409537
KEYWORDS    EST.
SOURCE      Lactuca sativa.
  ORGANISM  Lactuca sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE   1  (bases 1 to 629)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QG_CA_Contig1419, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QGF19  row: D  column: 10.
FEATURES             Location/Qualifiers
     source          1. .629
                     /organism="Lactuca sativa"
                     /cultivar="L.serriola"
                     /db_xref="taxon:4236"
                     /clone="QGF19D10"
                     /clone_lib="QG_EFGHJ lettuce serriola"
                     /lab_host="E.coli"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 10 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB-QG_EFGHJ lettuce serriola
                     TAG_TISSUE-flowers post-fertilized
                     TAG_SEQ-TGCCATCGGG"
BASE COUNT      138 a    137 c    183 g    170 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              150          Length:         629
Score:                  41.00        Matches:        7
Percent Similarity:     81.82%       Conservative:   2
Best Local Similarity:  63.64%       Mismatches:     2
Query Match:            87.23%       Indels:         0
DB:                     14           Gaps:           0

US-09-823-649A-2 (1-11) x BQ990012 (1-629)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   :::||||||||||:::
Db    471 CTTTCAGCAGAACTTTCCGTTCCGTATGAAGAT 439

RESULT 8
BQ856219
LOCUS       BQ856219                728 bp    mRNA    linear   EST 14-AUG-2002
DEFINITION  QGB28O14.yg.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone
            QGB28O14, mRNA sequence.
ACCESSION   BQ856219
VERSION     BQ856219.1  GI:22241684
KEYWORDS    EST.
SOURCE      Lactuca sativa.
  ORGANISM  Lactuca sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE   1  (bases 1 to 728)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QG_CA_Contig1419, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QGB28  row: O  column: 14.
FEATURES             Location/Qualifiers
     source          1. .728
                     /organism="Lactuca sativa"
                     /cultivar="Salinas"
                     /db_xref="taxon:4236"
                     /clone="QGB28O14"
                     /clone_lib="QG_ABCDI lettuce salinas"
```

/lab_host="E.coli"
/note="Vector: pBRcDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/ TAG_LIB=QG_ABCDI lettuce salinas TAG_TISSUE=chemical induction TAG_SEQ=TGTAGCCGGC"

BASE COUNT 203 a 203 c 153 g 169 t
ORIGIN

Alignment Scores:

Pred. No.:	182	Length:	728
Score:	41.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	87.23%	Indels:	0
DB:	14	Gaps:	0

US-09-823-649A-2 (1-11) x BQ856219 (1-728)

```
Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   :::||||||||||:::
Db     299 CTTTCAGCAGAACTTTCCGTTCCGTATGAAGAT 331
```

RESULT 9
BM411779
LOCUS BM411779 768 bp mRNA linear EST 22-JAN-2002
DEFINITION EST586106 tomato breaker fruit Lycopersicon esculentum cDNA clone cLEG57L16 5' end, mRNA sequence.
ACCESSION BM411779
VERSION BM411779.1 GI:18263409
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 768)
AUTHORS Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai,J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)
JOURNAL Unpublished (2002)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics Institute
Seq primer: T3.
FEATURES Location/Qualifiers
source 1..768
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG57L16"
/clone_lib="tomato breaker fruit"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCUadapt; Site_1: EcoR1; Site_2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 209 a 181 c 173 g 205 t
ORIGIN

Alignment Scores:

Pred. No.:	195	Length:	768
Score:	41.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	87.23%	Indels:	0
DB:	13	Gaps:	0

US-09-823-649A-2 (1-11) x BM411779 (1-768)

```
Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   ||||||   :::||||||||||:::
Db     261 CTCTCGGCTGAGCTCTCAGTTCCGTATGAGCAT 293
```

RESULT 10
BQ404238
LOCUS BQ404238 658 bp mRNA linear EST 22-MAY-2002
DEFINITION GA__Ed0068A08f Gossypium arboreum 7-10 dpa fiber library Gossypium arboreum cDNA clone GA__Ed0068A08f, mRNA sequence.
ACCESSION BQ404238
VERSION BQ404238.1 GI:21091925
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE 1 (bases 1 to 658)
AUTHORS Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber
JOURNAL Unpublished (2000)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total High Quality bases - 607
Seq primer: TAATACGACTCACTATAGGG
High quality sequence stop: 656.
FEATURES Location/Qualifiers
source 1..658
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA__Ed0068A08f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/note="Vector: pBK-CMV; Site_1: EcoR1; Site_2: XhoI"
BASE COUNT 170 a 164 c 122 g 201 t 1 others
ORIGIN

Alignment Scores:

Pred. No.:	256	Length:	658
Score:	40.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	85.11%	Indels:	0
DB:	14	Gaps:	0

US-09-823-649A-2 (1-11) x BQ404238 (1-658)

```
Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   :::||||||||||:::
Db    374 CTTTCCACTGAACTCTTTGTTCCGTATGAAGAT 406

RESULT 11
AV415207
LOCUS       AV415207                183 bp    mRNA    linear   EST 23-MAY-2000
DEFINITION  AV415207 Lotus japonicus young plants (two-week old) Lotus
            japonicus cDNA clone MWM104d04_r 5', mRNA sequence.
ACCESSION   AV415207
VERSION     AV415207.1  GI:7744383
KEYWORDS    EST.
SOURCE      Lotus japonicus.
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1  (bases 1 to 183)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
  JOURNAL   DNA Res. 7 (2), 127-130 (2000)
  MEDLINE   20277479
COMMENT     Contact: Yasukazu Nakamura
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .183
                     /organism="Lotus japonicus"
                     /db_xref="taxon:34305"
                     /clone="MWM104d04_r"
                     /clone_lib="Lotus japonicus young plants (two-week old)"
                     /dev_stage="young plants (two-week old)"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; isolate-Miyakojima MG-20"
BASE COUNT       33 a     64 c     44 g     42 t
ORIGIN

Alignment Scores:
Pred. No.:              125          Length:         183
Score:                  38.00        Matches:        7
Percent Similarity:     81.82%       Conservative:   2
Best Local Similarity:  63.64%       Mismatches:     2
Query Match:            80.85%       Indels:         0
DB:                     10           Gaps:           0

US-09-823-649A-2 (1-11) x AV415207 (1-183)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   ||||||   :::||||||:::|||
Db    117 CTCTCCGGCGAGCTCTGTGTTCCGTATGATGAA 149

RESULT 12
AV035497/c
LOCUS       AV035497                249 bp    mRNA    linear   EST 22-NOV-1999
DEFINITION  AV035497 Mus musculus adult C57BL/6J placenta Mus musculus cDNA
            clone 1600015A21, mRNA sequence.
ACCESSION   AV035497
VERSION     AV035497.1  GI:4855162
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 249)
  AUTHORS   Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
            Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T., Hara
            ,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M., Kawai,J.,
            Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H., Owa,C.,
            Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara
            ,Y., Suzuki,H., Suzuki,H., Tateno,M., Tomaru,Y., Tominaga,N.,
            Watanabe,S., Yagame,M., Yamamura,T., Yokota,T., Yoshino,M.,
            Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Chie Owa
            Genome Science Laboratory
            RIKEN
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-9145
            Fax: 81-298-36-9098
            Email: genome-res@rtc.riken.go.jp
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length cDNA
            (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
            Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
            Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .249
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="1600015A21"
                     /clone_lib="Mus musculus adult C57BL/6J placenta"
                     /sex="female"
                     /tissue_type="placenta"
                     /dev_stage="adult"
                     /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
BASE COUNT       69 a     40 c     40 g     100 t
ORIGIN

Alignment Scores:
Pred. No.:              186          Length:         249
Score:                  38.00        Matches:        7
Percent Similarity:     81.82%       Conservative:   2
Best Local Similarity:  63.64%       Mismatches:     2
Query Match:            80.85%       Indels:         0
DB:                     9            Gaps:           0

US-09-823-649A-2 (1-11) x AV035497 (1-249)

Qy      1 LeuSer***GluLeu***IleProTyrGluGlu 11
          ||||||   :::|||   :::||||||||||||
Db     76 TTGTCTAGGCAGCTGATCTTGCCATATGAAGAA 44

RESULT 13
AV428667
LOCUS       AV428667                411 bp    mRNA    linear   EST 02-MAY-2000
DEFINITION  AV428667 Lotus japonicus young plants (two-week old) Lotus
            japonicus cDNA clone MWL050g03_r 5', mRNA sequence.
ACCESSION   AV428667
VERSION     AV428667.1  GI:7678049
KEYWORDS    EST.
SOURCE      Lotus japonicus.
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
```

```
  TITLE     Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
  JOURNAL   DNA Res. 7 (2), 127-130 (2000)
  MEDLINE   20277479
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .411
                     /organism="Lotus japonicus"
                     /db_xref="taxon:34305"
                     /clone="MWL050g03_r"
                     /clone_lib="Lotus japonicus young plants (two-week old)"
                     /dev_stage="young plants (two-week old)"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; isolate=Miyakojima MG-20"
BASE COUNT       86 a    143 c     90 g     90 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              358          Length:          411
Score:                  38.00        Matches:         7
Percent Similarity:     81.82%       Conservative:    2
Best Local Similarity:  63.64%       Mismatches:      2
Query Match:            80.85%       Indels:          0
DB:                     10           Gaps:            0

US-09-823-649A-2 (1-11) x AV428667 (1-411)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   |||||||.. *::|||||||:::|||
Db     374 CTCTCCGGCGAGCTCTGTGTNCCGTATGATGAA 406

RESULT 14
BF758509/c
LOCUS       BF758509                416 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  MR4-CT0538-141100-104-b08 CT0538 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF758509
VERSION     BF758509.1  GI:12106409
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR4&t2=MR4-CT0538-
            141100-104-b08&t3=2000-11-14&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 30
            High quality sequence stop: 88.
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="CT0538"
                     /dev_stage="Adult"
                     /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT      129 a    106 c     85 g     96 t
ORIGIN

Alignment Scores:
Pred. No.:              363          Length:          416
Score:                  38.00        Matches:         8
Percent Similarity:     72.73%       Conservative:    0
Best Local Similarity:  72.73%       Mismatches:      3
Query Match:            80.85%       Indels:          0
DB:                     12           Gaps:            0

US-09-823-649A-2 (1-11) x BF758509 (1-416)

Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
           ||||||   |||       ||||||||||||||
Db     182 TTGTCTGGTGAGCGAGCTATCCCTTATGAGGAG 150

RESULT 15
AV415722
LOCUS       AV415722                426 bp    mRNA    linear   EST 23-MAY-2000
DEFINITION  AV415722 Lotus japonicus young plants (two-week old) Lotus
            japonicus cDNA clone MWM115e12_r 5', mRNA sequence.
ACCESSION   AV415722
VERSION     AV415722.1  GI:7744898
KEYWORDS    EST.
SOURCE      Lotus japonicus.
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Generation of 7137 non-redundant expressed sequence tags from a
            legume, Lotus japonicus
  JOURNAL   DNA Res. 7 (2), 127-130 (2000)
  MEDLINE   20277479
COMMENT     Contact: Yasukazu Nakamura
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: ynakamu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Lotus japonicus"
                     /db_xref="taxon:34305"
                     /clone="MWM115e12_r"
                     /clone_lib="Lotus japonicus young plants (two-week old)"
                     /dev_stage="young plants (two-week old)"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI; isolate=Miyakojima MG-20"
BASE COUNT       90 a    147 c     95 g     94 t
ORIGIN

Alignment Scores:
Pred. No.:              375          Length:          426
Score:                  38.00        Matches:         7
Percent Similarity:     81.82%       Conservative:    2
Best Local Similarity:  63.64%       Mismatches:      2
Query Match:            80.85%       Indels:          0
DB:                     10           Gaps:            0
```

US-09-823-649A-2 (1-11) x AV415722 (1-426)

```
Qy       1 LeuSer***GluLeu***IleProTyrGluGlu 11
             ||||||   ||||||   :::||||||:::|||
Db     375 CTCTCCGGCGAGCTCTGTGTTCCGTATGATGAA 407
```

Search completed: January 21, 2003, 12:24:13
Job time : 1171.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds
(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649A-6
Perfect score: 49
Sequence: 1 LSKRIGLSVSE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24626/app_query.fasta_1.1393
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649_@CGN_1_1_77_@runat_21012003_093151_24626
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	39	79.6	471	10	US-09-864-761-62	Sequence 62, Appl
c	2	39	79.6	668	10	US-09-864-761-16902	Sequence 16902, A
	3	37	75.5	1314	10	US-09-974-300-1325	Sequence 1325, Ap
	4	36	73.5	540	9	US-09-796-692-7302	Sequence 7302, Ap
	5	36	73.5	1024	10	US-09-925-301-15	Sequence 15, Appl
	6	35	71.4	193	10	US-09-864-761-26194	Sequence 26194, A
c	7	35	71.4	300	10	US-09-294-093B-3331	Sequence 3331, Ap
	8	35	71.4	568	10	US-09-864-761-9842	Sequence 9842, Ap
c	9	35	71.4	10885	10	US-09-764-877-3843	Sequence 3843, Ap
	10	35	71.4	126512	10	US-09-804-474A-3	Sequence 3, Appli
	11	35	71.4	155074	9	US-10-026-188-6	Sequence 6, Appli
	12	35	71.4	397658	10	US-09-813-320-3	Sequence 3, Appli
	13	34	69.4	431	10	US-09-954-456-1202	Sequence 1202, Ap
	14	34	69.4	431	10	US-09-880-107-709	Sequence 709, App
	15	34	69.4	462	10	US-09-864-761-5243	Sequence 5243, Ap
	16	34	69.4	471	10	US-09-604-287A-222	Sequence 222, App
	17	34	69.4	471	10	US-09-339-338-222	Sequence 222, App
	18	34	69.4	471	12	US-10-007-805-222	Sequence 222, App
c	19	34	69.4	614	10	US-09-925-300-653	Sequence 653, App
c	20	34	69.4	1083	10	US-09-974-300-2681	Sequence 2681, Ap
c	21	34	69.4	1491	10	US-09-808-483-13	Sequence 13, Appl
c	22	34	69.4	1833	10	US-09-815-242-6921	Sequence 6921, Ap
	23	34	69.4	2176	8	US-08-808-031A-1	Sequence 1, Appli
	24	34	69.4	2729	10	US-09-822-849A-529	Sequence 529, App
	25	34	69.4	2751	9	US-09-938-842A-1105	Sequence 1105, Ap
	26	34	69.4	25377	9	US-10-061-119-4	Sequence 4, Appli
	27	34	69.4	52354	10	US-09-742-311-3	Sequence 3, Appli
c	28	33	67.3	167	10	US-09-864-761-17003	Sequence 17003, A
c	29	33	67.3	219	10	US-09-560-863-409	Sequence 409, App
c	30	33	67.3	360	10	US-09-864-761-178	Sequence 178, App
c	31	33	67.3	477	10	US-09-815-242-3961	Sequence 3961, Ap
	32	33	67.3	480	10	US-09-764-887-121	Sequence 121, App
	33	33	67.3	507	9	US-10-046-935-494	Sequence 494, App
	34	33	67.3	507	9	US-09-878-178-494	Sequence 494, App
c	35	33	67.3	587	10	US-09-864-761-7966	Sequence 7966, Ap
c	36	33	67.3	754	10	US-09-910-943-330	Sequence 330, App
c	37	33	67.3	774	10	US-09-815-242-6584	Sequence 6584, Ap
	38	33	67.3	1081	10	US-09-070-927A-773	Sequence 773, App
c	39	33	67.3	1347	9	US-09-938-842A-268	Sequence 268, App
c	40	33	67.3	1970	10	US-09-864-761-4957	Sequence 4957, Ap
	41	33	67.3	2000	9	US-09-938-842A-3057	Sequence 3057, Ap
c	42	33	67.3	2769	9	US-09-911-176B-1	Sequence 1, Appli
	43	33	67.3	3072	10	US-09-981-900B-8	Sequence 8, Appli
	44	33	67.3	3209	10	US-09-220-091-4	Sequence 4, Appli
c	45	33	67.3	3723	12	US-10-001-215-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-864-761-62/c
; Sequence 62, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 62
;   LENGTH: 471
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005301.16
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
US-09-864-761-62

Alignment Scores:
Pred. No.:                4.39          Length:          471
Score:                    39.00         Matches:         8
Percent Similarity:       90.91%        Conservative:    2
Best Local Similarity:    72.73%        Mismatches:      1
Query Match:              79.59%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-864-761-62 (1-471)

Qy       1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
           |||:::|||||:::|||   ||||||||||||
Db     241 TTAACAAAACGGGTAGGCTCTTCAGTTTCAGAG 209

RESULT 2
US-09-864-761-16902/c
; Sequence 16902, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16902
;   LENGTH: 668
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005301.16
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
;   OTHER INFORMATION: SWISSPROT HIT: P51811, EVALUE 5.00e-43
;   OTHER INFORMATION: NT HIT: gi|11418230, EVALUE 1.00e-100
;   OTHER INFORMATION: EST_HUMAN HIT: BE791300.1, EVALUE 2.50e+00
US-09-864-761-16902

Alignment Scores:
Pred. No.:                6.66          Length:          668
Score:                    39.00         Matches:         8
Percent Similarity:       90.91%        Conservative:    2
Best Local Similarity:    72.73%        Mismatches:      1
Query Match:              79.59%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-864-761-16902 (1-668)

Qy       1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
           |||:::|||||:::|||   ||||||||||||
Db     207 TTAACAAAACGGGTAGGCTCTTCAGTTTCAGAG 175

RESULT 3
US-09-974-300-1325
; Sequence 1325, Application US/09974300
```

```
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1325
;   LENGTH: 1314
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1314)
;   OTHER INFORMATION: n = A,T,C or G
US-09-974-300-1325

Alignment Scores:
Pred. No.:                41.2          Length:          1314
Score:                    37.00         Matches:         8
Percent Similarity:       90.91%        Conservative:    2
Best Local Similarity:    72.73%        Mismatches:      1
Query Match:              75.51%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-974-300-1325 (1-1314)

Qy       1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
           ||||||:::|||   |||||||||:::||||||
Db     577 CTGTCAGAGAGAGCGGGGCTTTCGCTTTCAGAG 609

RESULT 4
US-09-796-692-7302
; Sequence 7302, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;  TITLE OF INVENTION:  HEMATOLOGICAL MALIGNANCIES
;  FILE REFERENCE: 2077.001200
;  CURRENT APPLICATION NUMBER: US/09/796,692
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 60/186,126
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 60/190,479
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/200,545
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 60/200,303
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,779
;  PRIOR FILING DATE: 2000-04-28
;  PRIOR APPLICATION NUMBER: 60/200,999
;  PRIOR FILING DATE: 2000-05-01
;  PRIOR APPLICATION NUMBER: 60/202,084
;  PRIOR FILING DATE: 2000-05-04
;  PRIOR APPLICATION NUMBER: 60/206,201
;  PRIOR FILING DATE: 2000-05-22
;  PRIOR APPLICATION NUMBER: 60/218,950
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/222,903
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 60/223,416
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: 60/223,378
;  PRIOR FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 9597
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7302
;   LENGTH: 540
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (502)
;   OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7302

Alignment Scores:
Pred. No.:                23.8          Length:          540
Score:                    36.00         Matches:         8
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              73.47%        Indels:          0
DB:                       9             Gaps:            0

US-09-823-649A-6 (1-11) x US-09-796-692-7302 (1-540)

Qy       4 ArgIleGlyLeuSerValSerGlu 11
           ||||||||||||||||||||||||
Db     360 AGGATTGGCCTGTCAGTTTCAGAG 383

RESULT 5
US-09-925-301-15
; Sequence 15, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA106
;  CURRENT APPLICATION NUMBER: US/09/925,301
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05882
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1694
;  SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 1024
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (938)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1005)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1012)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (1019)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-15

Alignment Scores:
Pred. No.:                50.9          Length:          1024
Score:                    36.00         Matches:         8
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              73.47%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-925-301-15 (1-1024)
```

```
Qy      4 ArgIleGlyLeuSerValSerGlu 11
          |||||||||||||||||||||||
Db    444 AGGATTGGCCTGTCAGTTTCAGAG 467

RESULT 6
US-09-864-761-26194
; Sequence 26194, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26194
;   LENGTH: 193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL022312.7
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
;   OTHER INFORMATION: NT HIT: AJ277662.1, EVALUE 1.10e-01
;   OTHER INFORMATION: SWISSPROT HIT: P50551, EVALUE 9.90e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BE393892.1, EVALUE 1.00e-104
US-09-864-761-26194

Alignment Scores:
Pred. No.:              11.6          Length:          193
Score:                  35.00         Matches:         8
Percent Similarity:     90.91%        Conservative:    2
Best Local Similarity:  72.73%        Mismatches:      1
Query Match:            71.43%        Indels:          0
DB:                     10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-864-761-26194 (1-193)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          ||||||:::|||||||   |||||||:::|||||||
Db     89 CTCAGTCGTCGGATCCAGCTCTCATTGTCAGAG 121

RESULT 7
US-09-294-093B-3331/c
; Sequence 3331, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
;  APPLICANT:       Lalgudi, Raghunath, V.
;  APPLICANT:  Ito, Laura, Y.
;  APPLICANT:  Sherman, Bradley, K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
;  FILE REFERENCE:  PL-0009 US
;  CURRENT APPLICATION NUMBER: US/09/294,093B
;  CURRENT FILING DATE:  1999-04-16
;  PRIOR APPLICATION NUMBER:  60/082,567
;  PRIOR FILING DATE:  April 21, 1998
;  NUMBER OF SEQ ID NOS:      6207
;  SOFTWARE:       PERL Program
; SEQ ID NO 3331
;   LENGTH: 300
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20010051335A1 700379842H1
;   NAME/KEY: unsure
;   LOCATION: 93
;   OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-3331

Alignment Scores:
Pred. No.:              19.6          Length:          300
Score:                  35.00         Matches:         7
Percent Similarity:     100.00%       Conservative:    3
Best Local Similarity:  70.00%        Mismatches:      0
Query Match:            71.43%        Indels:          0
DB:                     10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-294-093B-3331 (1-300)

Qy      1 LeuSerLysArgIleGlyLeuSerValSer 10
          ||||||:::|||:::|||:::||||||||||
Db    210 CTCTCCAGACGCATGGGCGTCTCCGTCTCT 181

RESULT 8
US-09-864-761-9842
; Sequence 9842, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
```

```
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9842
;   LENGTH: 568
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL022312.7
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 1.4
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 2.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 1.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 1.5
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 2.6
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 1.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 1.4
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 2.4
US-09-864-761-9842

Alignment Scores:
Pred. No.:                41.9          Length:          568
Score:                    35.00         Matches:         8
Percent Similarity:       90.91%        Conservative:    2
Best Local Similarity:    72.73%        Mismatches:      1
Query Match:              71.43%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-864-761-9842 (1-568)

Qy       1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
           ||||||:::|||||||   |||||||:::|||||||
Db     290 CTCAGTCGTCGGATCCAGCTCTCATTGTCAGAG 322

RESULT 9
US-09-764-877-3843/c
; Sequence 3843, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3843
;   LENGTH: 10885
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-877-3843

Alignment Scores:
Pred. No.:                1.41e+03      Length:          10885
Score:                    35.00         Matches:         7
Percent Similarity:       90.00%        Conservative:    2
Best Local Similarity:    70.00%        Mismatches:      1
Query Match:              71.43%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-764-877-3843 (1-10885)

Qy       1 LeuSerLysArgIleGlyLeuSerValSer 10
           ||||||   |||||||||:::|||:::|||
Db    1141 CTTAGCAATAGAATTGGGGTCTCTATTTCT 1112

RESULT 10
US-09-804-474A-3
; Sequence 3, Application US/09804474A
; Patent No. US20020119518A1
; GENERAL INFORMATION:
;  APPLICANT: KODET, Stefan et al
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000891
;  CURRENT APPLICATION NUMBER: US/09/804,474A
;  CURRENT FILING DATE:  2001-03-13
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 126512
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(126512)
;   OTHER INFORMATION: n = A,T,C or G
US-09-804-474A-3

Alignment Scores:
Pred. No.:                2.57e+04      Length:          126512
Score:                    35.00         Matches:         7
Percent Similarity:       90.00%        Conservative:    2
Best Local Similarity:    70.00%        Mismatches:      1
Query Match:              71.43%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-804-474A-3 (1-126512)

Qy       2 SerLysArgIleGlyLeuSerValSerGlu 11
           :::|||||||||||||||   |||||||:::
Db 123466 AATAAAAGAATTGGGCTAGATGTCTCTGAC 123495

RESULT 11
US-10-026-188-6
; Sequence 6, Application US/10026188
; Patent No. US20020164645A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Zhang, Yifeng
```

```
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
;  TITLE OF INVENTION:  Ion Channel
;  FILE REFERENCE: 02307E-114910US
;  CURRENT APPLICATION NUMBER: US/10/026,188
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: US 60/259,379
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 155074
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human genomic region containing ltrpc5
;   OTHER INFORMATION: (Human Chromosome 11p15.5 PAC clone pDJ915f1
;   OTHER INFORMATION: containing KvLQT1 gene)
US-10-026-188-6

Alignment Scores:
Pred. No.:              3.26e+04        Length:         155074
Score:                  35.00           Matches:        7
Percent Similarity:     72.73%          Conservative:   1
Best Local Similarity:  63.64%          Mismatches:     3
Query Match:            71.43%          Indels:         0
DB:                     9               Gaps:           0

US-09-823-649A-6 (1-11) x US-10-026-188-6 (1-155074)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          |||   ||||||:::|||||||   |||   |||
Db  28786 CTAGGGAAAAGATTGGGCTTGGGGGTTGGGGAG 28818

RESULT 12
US-09-813-320-3
; Sequence 3, Application US/09813320
; Patent No. US20020142378A1
; GENERAL INFORMATION:
;  APPLICANT: ZHANG, Hongyu et al.
;  TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL001172
;  CURRENT APPLICATION NUMBER: US/09/813,320
;  CURRENT FILING DATE:  2001-03-21
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 397658
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(397658)
;   OTHER INFORMATION: n = A,T,C or G
US-09-813-320-3

Alignment Scores:
Pred. No.:              9.57e+04        Length:         397658
Score:                  35.00           Matches:        6
Percent Similarity:     90.91%          Conservative:   4
Best Local Similarity:  54.55%          Mismatches:     1
Query Match:            71.43%          Indels:         0
DB:                     10              Gaps:           0

US-09-823-649A-6 (1-11) x US-09-813-320-3 (1-397658)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          |||   |||:::::|||:::|||:::||||||
Db 101718 CTCCCAAAGAAGATGGGCATCAGTATTTCAGAA 101750

RESULT 13
US-09-954-456-1202
; Sequence 1202, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1202
;   LENGTH: 431
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-1202

Alignment Scores:
Pred. No.:              50.2            Length:         431
Score:                  34.00           Matches:        6
Percent Similarity:     100.00%         Conservative:   4
Best Local Similarity:  60.00%          Mismatches:     0
Query Match:            69.39%          Indels:         0
DB:                     10              Gaps:           0

US-09-823-649A-6 (1-11) x US-09-954-456-1202 (1-431)

Qy      1 LeuSerLysArgIleGlyLeuSerValSer 10
          :::|||:::|||:::|||||||:::|||||
Db    263 GTCTCAAGGAGGGTGGGGCTTGCAGTCAGT 292

RESULT 14
US-09-880-107-709
; Sequence 709, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 709
;   LENGTH: 431
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA292440
US-09-880-107-709

Alignment Scores:
Pred. No.:              50.2          Length:          431
Score:                  34.00         Matches:         6
Percent Similarity:     100.00%       Conservative:    4
Best Local Similarity:  60.00%        Mismatches:      0
Query Match:            69.39%        Indels:          0
DB:                     10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-880-107-709 (1-431)

Qy       1 LeuSerLysArgIleGlyLeuSerValSer 10
           :::|||:::|||:::|||||||:::|||||
Db     263 GTCTCAAGGAGGGTGGGGCTTGCAGTCAGT 292

RESULT 15
US-09-864-761-5243
; Sequence 5243, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5243
;   LENGTH: 462
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004954.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =  18
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL =   6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   6
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.5
US-09-864-761-5243

Alignment Scores:
Pred. No.:              54.5          Length:          462
Score:                  34.00         Matches:         7
Percent Similarity:     88.89%        Conservative:    1
Best Local Similarity:  77.78%        Mismatches:      1
Query Match:            69.39%        Indels:          0
DB:                     10            Gaps:            0

US-09-823-649A-6 (1-11) x US-09-864-761-5243 (1-462)

Qy       1 LeuSerLysArgIleGlyLeuSerVal 9
           |||   ||||||:::|||||||||||
Db     187 TTAGAGAAAAGGCTGGGGCTTTCTGTC 213


Search completed: January 21, 2003, 10:08:27
Job time : 66.5714 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds
(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649A-6
Perfect score: 49
Sequence: 1 LSKRIGLSVSE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24616/app_query.fasta_1.1393
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649_@CGN_1_1_2820_@runat_21012003_093151_24616 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	41	83.7	429	9	AU228362	AU228362 AU228362
c 2	40	81.6	509	17	AZ929376	AZ929376 479.dif25
c 3	40	81.6	1117	17	CNS06M4N	AL405021 T7 end of
c 4	39	79.6	383	10	AW869477	AW869477 MR3-SN006
5	39	79.6	691	17	AG085155	AG085155 Pan trog1
6	38	77.6	413	12	BG096912	BG096912 EST461431
7	38	77.6	425	14	BM871118	BM871118 mgns013xA
8	38	77.6	442	10	AW097231	AW097231 rs41d05.y
9	38	77.6	495	17	BH828859	BH828859 BACPP28-G
10	38	77.6	545	12	BF049437	BF049437 db81d03.x
c 11	38	77.6	613	10	AW958920	AW958920 EST370990
c 12	38	77.6	666	13	BJ165314	BJ165314 BJ165314
c 13	38	77.6	812	14	BQ506021	BQ506021 EST613436
c 14	37	75.5	313	9	AA572148	AA572148 v150e12.r
c 15	37	75.5	350	9	AU112697	AU112697 AU112697
c 16	37	75.5	368	13	BJ389576	BJ389576 BJ389576
c 17	37	75.5	376	14	C10047	C10047 C10047 Yuji
c 18	37	75.5	378	9	AV188048	AV188048 AV188048
19	37	75.5	413	17	B77147	B77147 T31B4TF TAM
c 20	37	75.5	487	13	BJ325691	BJ325691 BJ325691
21	37	75.5	497	12	BG018916	BG018916 dab02d04.
c 22	37	75.5	502	13	BJ362913	BJ362913 BJ362913
c 23	37	75.5	515	17	AQ788974	AQ788974 HS_3044_B
c 24	37	75.5	526	14	BM953398	BM953398 952061H02
c 25	37	75.5	540	17	AZ258482	AZ258482 RPCI-23-1
c 26	37	75.5	541	12	BF246262	BF246262 601853864
c 27	37	75.5	554	13	BJ363889	BJ363889 BJ363889
c 28	37	75.5	556	13	BJ410239	BJ410239 BJ410239
29	37	75.5	576	13	BJ093481	BJ093481 BJ093481
30	37	75.5	596	13	BJ142376	BJ142376 BJ142376
31	37	75.5	599	17	AZ130889	AZ130889 OSJNBb007
32	37	75.5	638	9	AL681825	AL681825 AL681825
33	37	75.5	654	13	BJ070414	BJ070414 BJ070414
34	37	75.5	662	13	BJ075163	BJ075163 BJ075163
35	37	75.5	668	13	BJ346356	BJ346356 BJ346356
c 36	37	75.5	705	17	AG105753	AG105753 Pan trog1
37	37	75.5	707	14	BQ181293	BQ181293 UI-H-EU0-
c 38	37	75.5	723	9	AI936015	AI936015 wn79q12.x
39	37	75.5	728	12	BG494983	BG494983 602541612
40	37	75.5	751	12	BF615515	BF615515 de08b04.y
c 41	37	75.5	760	13	BI903478	BI903478 603167007
42	37	75.5	779	17	BH305311	BH305311 CH230-192
c 43	37	75.5	797	17	BH094735	BH094735 RPCI-24-2
44	37	75.5	805	12	BG722042	BG722042 602698632
c 45	37	75.5	809	12	BF541083	BF541083 602068667

ALIGNMENTS

RESULT 1
AU228362/c
LOCUS AU228362 429 bp mRNA linear EST 23-APR-2002
DEFINITION AU228362 RAFL16 Arabidopsis thaliana cDNA clone RAFL16-13-I07 3', mRNA sequence.
ACCESSION AU228362
VERSION AU228362.1 GI:19743009
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 429)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,

```
            Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
            Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu
            ,M., Hayashizaki,Y. and Shinozaki,K.
  TITLE     Large scale analysis of Arabidopsis full-length cDNA
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Croup
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998).cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
            al., submitted for publication) digested with BamHI and SalI. This
            clone is in a modified pBluescript vector. Please visit our web
            site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
            details.
FEATURES             Location/Qualifiers
     source          1. .429
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="RAFL16-13-107"
                     /clone_lib="RAFL16"
                     /lab_host="DH10B"
                     /note="Site_1: BamHI; Site_2: SalI; dark-grown"
BASE COUNT      156 a     86 c     74 g    113 t
ORIGIN

Alignment Scores:
Pred. No.:              153           Length:         429
Score:                  41.00         Matches:        9
Percent Similarity:     90.91%        Conservative:   1
Best Local Similarity:  81.82%        Mismatches:     1
Query Match:            83.67%        Indels:         0
DB:                     9             Gaps:           0

US-09-823-649A-6 (1-11) x AU228362 (1-429)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          |||||||||:::|||||||||   |||||||||
Db    123 TTCTCCAAACAAATCGGTTTCAAGGTTTCAGAA 91

RESULT 2
AZ929376/c
LOCUS       AZ929376                509 bp    DNA     linear   GSS 01-APR-2001
DEFINITION  479.dif25e12.s1 Saccharomyces kluyveri Saccharomyces kluyveri
            genomic clone 479.dif25e12.s1, DNA sequence.
ACCESSION   AZ929376
VERSION     AZ929376.1  GI:13500283
KEYWORDS    GSS.
SOURCE      Saccharomyces kluyveri.
  ORGANISM  Saccharomyces kluyveri
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 509)
  AUTHORS   Cliften,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish
            ,W.R., Waterston,R.H. and Johnston,M.
  TITLE     Surveying Saccharomyces genomes to identify functional elements by
            comparative DNA sequence analysis
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Johnston M
            Department of Genetics
            Washington University Medical School
            Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
            Tel: 314 362 2735
            Fax: 314 362 7855
            Email: mj@genetics.wustl.edu
            Class: random plasmid subclone.
FEATURES             Location/Qualifiers
     source          1. .509
                     /organism="Saccharomyces kluyveri"
                     /strain="NRRL Y-12651 (CBS 3082)"
                     /db_xref="taxon:4934"
                     /clone="479.dif25e12.s1"
                     /clone_lib="Saccharomyces kluyveri"
                     /note="Random genomic sequence "
BASE COUNT      154 a     103 c     84 g    168 t
ORIGIN

Alignment Scores:
Pred. No.:              309           Length:         509
Score:                  40.00         Matches:        9
Percent Similarity:     81.82%        Conservative:   0
Best Local Similarity:  81.82%        Mismatches:     2
Query Match:            81.63%        Indels:         0
DB:                     17            Gaps:           0

US-09-823-649A-6 (1-11) x AZ929376 (1-509)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          |||||||||||||||||||||   |||   |||
Db    101 TTGTCAAAAAGGATTGGGCTACGTGTAAAAGAG 69

RESULT 3
CNS06M4N/c
LOCUS       CNS06M4N                1117 bp    DNA     linear   GSS 17-JUN-2001
DEFINITION  T7 end of clone AU0AA002A02 of library AU0AA from strain CBS 3082
            of Saccharomyces kluyveri, genomic survey sequence.
ACCESSION   AL405021
VERSION     AL405021.1  GI:12167229
KEYWORDS    GSS.
SOURCE      Saccharomyces kluyveri.
  ORGANISM  Saccharomyces kluyveri
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 1117)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 1117)
  AUTHORS   Neuveglise,C., Bon,E., Lepingle,A., Wincker,P., Artiguenave,F.,
            Gaillardin,C. and Casaregola,S.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 9.
            Saccharomyces kluyveri
  JOURNAL   FEBS Lett. 487 (1), 56-60 (2000)
  MEDLINE   20584719
   PUBMED   11152884
REFERENCE   3  (bases 1 to 1117)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiquus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .1117
                     /organism="Saccharomyces kluyveri"
```

```
                     /strain="CBS 3082"
                     /db_xref="taxon:4934"
                     /clone="AU0AA002A02"
                     /clone_lib="AU0AA"
                     /note="end : T7"
     misc_feature    <31. .>1042
                     /note="similar to Saccharomyces cerevisiae ORF YHR099w [
                     TRA1 ; strong similarity to human TRRAP protein ]
                     1 putative frameshift(s)"
                     /evidence=not_experimental
BASE COUNT      332 a    230 c    186 g    367 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              889           Length:        1117
Score:                  40.00         Matches:       9
Percent Similarity:     81.82%        Conservative:  0
Best Local Similarity:  81.82%        Mismatches:    2
Query Match:            81.63%        Indels:        0
DB:                     17            Gaps:          0

US-09-823-649A-6 (1-11) x CNS06M4N (1-1117)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          ||||||||||||||||||||   |||   |||
Db    482 TTGTCAAAAAGGATTGGGCTACGTGTAAAAGAG 450

RESULT 4
AW869477/c
LOCUS       AW869477                383 bp    mRNA    linear   EST 22-MAY-2000
DEFINITION  MR3-SN0068-070500-002-b10 SN0068 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW869477
VERSION     AW869477.1  GI:8003530
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 383)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=MR3-SN0068-070
            500-002-b10&t3=2000-05-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 16
            High quality sequence stop: 197.
FEATURES            Location/Qualifiers
     source         1. .383
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_lib="SN0068"
                    /dev_stage="Adult"
                    /note="Organ: stomach_normal; Vector: puc18; Site_1: SmaI;
                    Site_2: SmaI; A mini-library was made by cloning products
                    derived from  ORESTES PCR (U.S. Letters Patent application
                    No. 196,716 -  Ludwig Institute for Cancer Research)
                    profiles into  the pUC 18 vector. Reverse transcription of
                    tissue mRNA and cDNA amplification were performed under
                    low stringency conditions."
BASE COUNT      134 a    102 c     87 g     60 t
ORIGIN

Alignment Scores:
Pred. No.:              339           Length:        383
Score:                  39.00         Matches:       8
Percent Similarity:     100.00%       Conservative:  3
Best Local Similarity:  72.73%        Mismatches:    0
Query Match:            79.59%        Indels:        0
DB:                     10            Gaps:          0

US-09-823-649A-6 (1-11) x AW869477 (1-383)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          ||||||:::|||||||||||||:::|||:::|||
Db    140 CTGTCTGAGAGAATTGGTCTTACGGTGACCGAG 108

RESULT 5
AG085155
LOCUS       AG085155                691 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-083B22.F, genomic survey sequence.
ACCESSION   AG085155
VERSION     AG085155.1  GI:16636957
KEYWORDS    GSS.
SOURCE      Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
            BAC Library clone:PTB-083B22.F.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 691)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-cho,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: -21M13
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES            Location/Qualifiers
     source         1. .691
                    /organism="Pan troglodytes"
                    /db_xref="taxon:9598"
                    /clone="PTB-083B22.F"
                    /sex="male"
                    /cell_type="lymphoblast"
                    /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT      230 a    151 c    127 g    181 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              750           Length:        691
Score:                  39.00         Matches:       8
Percent Similarity:     90.91%        Conservative:  2
Best Local Similarity:  72.73%        Mismatches:    1
Query Match:            79.59%        Indels:        0
```

```
DB:                    17            Gaps:          0

US-09-823-649A-6 (1-11) x AG085155 (1-691)

Qy       1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
           |||:::|||||:::|||   ||||||||||
Db     132 TTAACAAAAAGGGTAGGCTCTTCAGTTTCAGAG 164

RESULT 6
BG096912
LOCUS       BG096912                413 bp    mRNA    linear   EST 29-JAN-2001
DEFINITION  EST461431 potato leaves and petioles Solanum tuberosum cDNA clone
            cSTB45C24 5' sequence, mRNA sequence.
ACCESSION   BG096912
VERSION     BG096912.1  GI:12586947
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I  Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
            Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
            ,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
  TITLE     Generation of ESTs from potato leaves and petioles
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cathy Ronning
            The Institute for Genomic Research
            For clone info: please contact Research Genetics, Libraries
            Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES             Location/Qualifiers
     source          1. .413
                     /organism="Solanum tuberosum"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="cSTB45C24"
                     /clone_lib="potato leaves and petioles"
                     /tissue_type="leaflets and petioles"
                     /dev_stage="8 weeks old plants"
                     /lab_host="SOLR"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Tissue was supplied by Dr. Fry (Cornell University).
                     Leaflets and petioles were isolated from 8 week old
                     greenhouse grown plants. The plants were watered and
                     fertilized freely. The tissue was immediately frozen in
                     liquid nitrogen."
BASE COUNT      151 a     59 c     83 g    120 t
ORIGIN

Alignment Scores:
Pred. No.:              604           Length:        413
Score:                  38.00         Matches:       7
Percent Similarity:     100.00%       Conservative:  2
Best Local Similarity:  77.78%        Mismatches:    0
Query Match:            77.55%        Indels:        0
DB:                     12            Gaps:          0

US-09-823-649A-6 (1-11) x BG096912 (1-413)

Qy       2 SerLysArgIleGlyLeuSerValSer 10
           ||||||||||:::||||||||||:::|||
Db     253 TCCAAAAGAGTTGGGCTTTCAATCAGT 279

RESULT 7
BM871118
LOCUS       BM871118                425 bp    mRNA    linear   EST 07-MAR-2002
DEFINITION  mgns013xA02f.b Magnaporthe grisea NS Uni-Zap XR Library Magnaporthe
            grisea cDNA clone mgns013xA02 5', mRNA sequence.
ACCESSION   BM871118
VERSION     BM871118.1  GI:19238800
KEYWORDS    EST.
SOURCE      Magnaporthe grisea.
  ORGANISM  Magnaporthe grisea
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
REFERENCE   1  (bases 1 to 425)
  AUTHORS   Ebbole,D.J., Yuan,J., Thomas,T.L., Bobrowicz,P. and Dean,R.A.
  TITLE     Expressed sequence tags from the rice blast fungus, Magnaporthe
            grisea
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Ebbole DJ
            Department of Plant Pathology & Microbiology
            Texas A&M University
            Peterson Bldg, MS2132, College Station, TX 77843-2132, USA
            Tel: 979 845 4831
            Fax: 979 845 6483
            Email: d-ebbole@tamu.edu
            Chromatogram file of this sequence is available, see contact person
            ;Best nr hit (Nov. 11, 2001) pir|T50972 probable zuotin [imported]
            - Neurospora crassa >gi|9. . . 148 4e-35
            PCR PRimers
            FORWARD: T3 primer
            BACKWARD: T7 primer
            Plate: mgns013  row: A  column: 02
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .425
                     /organism="Magnaporthe grisea"
                     /strain="Guy11"
                     /db_xref="taxon:148305"
                     /clone="mgns013xA02"
                     /clone_lib="Magnaporthe grisea NS Uni-Zap XR Library"
                     /sex="Mat1-2 hermaphrodite"
                     /cell_type="mycelium"
                     /note="Vector: pBluescriptSK-; Site_1: EcoRI; Site_2: XhoI
                     ; Unidirectional cloning. EcoRI side has T3 primer and
                     predominantly 5' reads. T7 primer on XhoI side of insert.
                     Nitrogen starvation library. Cells were inoculated into
                     minimal medium and grown for two days with shaking (150
                     rpm) at room temperature. Culture was harvested, blended,
                     inoculated into minimal medium as above for 24 h. Cells
                     were harvested, washed with water and inoculated into
                     minimal medium base lacking nitrogen source for 6 h.
                     Sequences were processed by phredPhrap 991019 and trimmed
                     according to phd files and for vector seqs."
BASE COUNT       91 a    128 c     90 g    116 t
ORIGIN

Alignment Scores:
Pred. No.:              628           Length:        425
Score:                  38.00         Matches:       8
Percent Similarity:     90.91%        Conservative:  2
Best Local Similarity:  72.73%        Mismatches:    1
Query Match:            77.55%        Indels:        0
DB:                     14            Gaps:          0

US-09-823-649A-6 (1-11) x BM871118 (1-425)

Qy       1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
           ||||||||||||:::||||||:::|||   |||
Db     120 CTAAGCAAAAGACTTGGCCTCGCCGTCCTTCAG 152

RESULT 8
AW097231
LOCUS       AW097231                442 bp    mRNA    linear   EST 10-MAY-2001
DEFINITION  rs41d05.y1 Sommer Pristionchus Pristionchus pacificus cDNA 5', mRNA
            sequence.
ACCESSION   AW097231
VERSION     AW097231.1  GI:6067542
KEYWORDS    EST.
SOURCE      Pristionchus pacificus.
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 442)
```

```
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Dr. Ralf Sommer DNA Sequencing by:
            Washington University Genome Sequencing Center
             Contact Dr. Ralf Sommer (ralf.sommer@tuebingen.mpg.de) for
            information about this clone.
            Putative full length read
            The vector to vector length is 738
            Seq primer: -40RP from Gibco
            High quality sequence stop: 404.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Pristionchus pacificus"
                     /strain="PS 312"
                     /db_xref="taxon:54126"
                     /clone_lib="Sommer Pristionchus"
                     /sex="predominantly hermaphroditic"
                     /dev_stage="mixed stages (embryo to adult)"
                     /lab_host="not applicable (host cell line)"
                     /note="Vector: Uni-ZAP XR Vector (Stratagene); Site_1: 5'
                     EcoRI; Site_2: 3' XhoI; 1st strand cDNA was primed with a
                     XhoI - oligo(dT) primer. Double-stranded cDNA was ligated
                     to EcoRI adaptors digested with XhoI and cloned into XhoI
                     and EcoRI sites. Primary complexity of the library was 10
                     in the 7th. The library went through one round of
                     amplification."
BASE COUNT      141 a     95 c    109 g     97 t
ORIGIN

Alignment Scores:
Pred. No.:                662           Length:         442
Score:                    38.00         Matches:        7
Percent Similarity:       90.91%        Conservative:   3
Best Local Similarity:    63.64%        Mismatches:     1
Query Match:              77.55%        Indels:         0
DB:                       10            Gaps:           0

US-09-823-649A-6 (1-11) x AW097231 (1-442)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          ||||||||||||:::|||:::   :::||||||
Db    164 TTGAGCAAGAGGCTCGGAATTGACTTAAGTGAA 196

RESULT 9
BH828859
LOCUS       BH828859                 495 bp    DNA     linear   GSS 20-MAY-2002
DEFINITION  BACPP28-G18.y Pristionchus pacificus BAC ends Pristionchus
            pacificus genomic, DNA sequence.
ACCESSION   BH828859
VERSION     BH828859.1  GI:21023822
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus.
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   Srinivasan,J., Sinz,W., Lanz,C., Brand,A., Nandakumar,R., Raddatz
            ,G., Witte,H., Keller,H., Kipping,I., Pires da Silva,A., Jesse,T.,
            Millare,J., de Both,M., Schuster,S.C. and Sommer,R.J.
  TITLE     A BAC-based genetic linkage map of the nematode Pristionchus
            pacificus
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .495
                     /organism="Pristionchus pacificus"
                     /strain="var. California"
                     /db_xref="taxon:54126"
                     /clone_lib="Pristionchus pacificus BAC ends"
BASE COUNT      142 a     85 c    105 g    163 t
ORIGIN

Alignment Scores:
Pred. No.:                771           Length:         495
Score:                    38.00         Matches:        7
Percent Similarity:       90.91%        Conservative:   3
Best Local Similarity:    63.64%        Mismatches:     1
Query Match:              77.55%        Indels:         0
DB:                       17            Gaps:           0

US-09-823-649A-6 (1-11) x BH828859 (1-495)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          ||||||||||||:::|||:::   :::||||||
Db    269 TTGACCAAGAGGCTCGGAATTCACTTAAGTGAG 301

RESULT 10
BF049437
LOCUS       BF049437                 545 bp    mRNA    linear   EST 11-OCT-2000
DEFINITION  db81d03.x1 Wellcome CRC pSK egg Xenopus laevis cDNA clone
            IMAGE:3379397 3' similar to SW:RL11_HUMAN P39026 60S RIBOSOMAL
            PROTEIN L11. [1] ;, mRNA sequence.
ACCESSION   BF049437
VERSION     BF049437.1  GI:10768022
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
            , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
            ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Xenopus EST project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: db81d03.y1
            Contact: Sandy Clifton, Ph.D.
            WashU Xenopus EST project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B.
            Gurdon (Wellcome/CRC Institute). DNA Sequencing by: Washington
            University Genome Sequencing Center
             Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 423.
FEATURES             Location/Qualifiers
     source          1. .545
                     /organism="Xenopus laevis"
```

```
                     /db_xref="taxon:8355"
                     /clone="IMAGE:3379397"
                     /clone_lib="Wellcome CRC pSK egg"
                     /tissue_type="egg"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pBluescript SK-; Site_1: NotI; Site_2:
                     EcoRI; cDNAs were oligo-dT primed and directionally
                     cloned. Library was constructed by N. Garrett, P. LeMaire,
                     A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute).
                     Note: This is a Xenopus Gene Collection (XGC) library."
BASE COUNT      129 a    153 c    105 g    158 t
ORIGIN

Alignment Scores:
Pred. No.:              877          Length:        545
Score:                  38.00        Matches:       8
Percent Similarity:     90.00%       Conservative:  1
Best Local Similarity:  80.00%       Mismatches:    1
Query Match:            77.55%       Indels:        0
DB:                     12           Gaps:          0

US-09-823-649A-6 (1-11) x BF049437 (1-545)

Qy      1 LeuSerLysArgIleGlyLeuSerValSer 10
          ||||||   ||||||||||||||||:::|||
Db    508 CTCAGCTCACGCATAGGGTTATCTATCTCA 537

RESULT 11
AW958920/c
LOCUS       AW958920                613 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST370990 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW958920
VERSION     AW958920.1  GI:8148604
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 613)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 127
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1. .613
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGE"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      175 a    130 c    158 g    149 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.03e+03     Length:        613
Score:                  38.00        Matches:       8
Percent Similarity:     90.91%       Conservative:  2
Best Local Similarity:  72.73%       Mismatches:    1
Query Match:            77.55%       Indels:        0
DB:                     10           Gaps:          0

US-09-823-649A-6 (1-11) x AW958920 (1-613)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          ||||||:::|||:::|||||||||   ||||||
Db    606 CTTTCAAGAAGATTGGGTCTAAGTCCTTCAGAA 574

RESULT 12
BJ165314/c
LOCUS       BJ165314                666 bp    mRNA    linear   EST 24-JAN-2002
DEFINITION  BJ165314 full length cDNA library, chloronemata and young
            gametophores Physcomitrella patens subsp. patens cDNA clone pph3n09
            3', mRNA sequence.
ACCESSION   BJ165314
VERSION     BJ165314.1  GI:18333300
KEYWORDS    EST.
SOURCE      Physcomitrella patens subsp. patens.
  ORGANISM  Physcomitrella patens subsp. patens
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE   1  (bases 1 to 666)
  AUTHORS   Fujita,T., Shin-i,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.
            , Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe
            ,M.
  TITLE     Comparison of the moss Physcomitrella patens genome with flowering
            plants genome
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            A backbone of the vector is basically from pBluescript(KS), that
            was in vivo excised from a modified lPS phage vector (Mo bi Tec,
            Germany). 5' end of the cDNA that was digested with XhoI was
            ligated to SalI site of the vector and the 3' end including polyA
            tail was ligated to BamHI site of the vector. cDNA instert could be
            amplified with conventional T7 and T3 primers. This full-length
            cDNA library was generated basically according to the method
            described in The Plant J 15, 707-720 (1998) Seki M. et al.
            Protonemata were blended by the POLYTRON, and then cultivated on
            the BCDATG medium for 13-14 days under the continuous light.
FEATURES             Location/Qualifiers
     source          1. .666
                     /organism="Physcomitrella patens subsp. patens"
                     /db_xref="taxon:145481"
                     /clone="pph3n09"
                     /clone_lib="full length cDNA library, chloronemata and
                     young gametophores"
                     /tissue_type="mixture of chloronemata and young
                     gametophores with 2 to 5 leaves"
BASE COUNT      150 a    208 c    141 g    167 t
ORIGIN

Alignment Scores:
Pred. No.:              1.15e+03     Length:        666
Score:                  38.00        Matches:       9
Percent Similarity:     90.91%       Conservative:  1
Best Local Similarity:  81.82%       Mismatches:    1
Query Match:            77.55%       Indels:        0
DB:                     13           Gaps:          0

US-09-823-649A-6 (1-11) x BJ165314 (1-666)

Qy      1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
          ||||||:::|||||||||||||   |||||||||
Db     94 TTGTCTCGGAGAATTGGGCTATGAGTCTCGGAA 62

RESULT 13
BQ506021/c
LOCUS       BQ506021                812 bp    mRNA    linear   EST 22-JUL-2002
DEFINITION  EST613436 Generation of a set of potato cDNA clones for microarray
            analyses mixed potato tissues Solanum tuberosum cDNA clone STMGH75
            3' end, mRNA sequence.
ACCESSION   BQ506021
```

```
VERSION     BQ506021.2  GI:21921918
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 812)
  AUTHORS   Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
  TITLE     Generation of a set of potato cDNA clones for microarray analyses
  JOURNAL   Unpublished (2002)
COMMENT     On Jun 10, 2002 this sequence version replaced gi:21364890.
            Other_ESTs: EST613435
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato@tigr.org
            This clone is available through the Research Genetics, contact the
            Research Genetics for further information 1-800-711-6195 or
            cdna@resgen.com
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .812
                     /organism="Solanum tuberosum"
                     /cultivar="Kennebec or Binjte"
                     /db_xref="taxon:4113"
                     /clone="STMGH75"
                     /clone_lib="Generation of a set of potato cDNA clones for
                     microarray analyses mixed potato tissues"
                     /tissue_type="mixed tissues"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Combination of untreated and Phytophthora
                     infestans-treated libraries of stolons, leaves, leaflets,
                     axillary buds of stem explants, petioles, germinating eyes
                     , tubers, or roots."
BASE COUNT      229 a    175 c    161 g    247 t
ORIGIN

Alignment Scores:
Pred. No.:              1.5e+03        Length:        812
Score:                  38.00          Matches:       8
Percent Similarity:     90.00%         Conservative:  1
Best Local Similarity:  80.00%         Mismatches:    1
Query Match:            77.55%         Indels:        0
DB:                     14             Gaps:          0

US-09-823-649A-6 (1-11) x BQ506021 (1-812)

Qy      1 LeuSerLysArgIleGlyLeuSerValSer 10
          |||   ||||||||||||:::|||||||||
Db    593 TTGCAGAAAAGGATAGGAATTTCAGTATCC 564

RESULT 14
AA572148/c
LOCUS       AA572148                313 bp    mRNA    linear   EST 27-AUG-1997
DEFINITION  vl50e12.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:975694 5', mRNA sequence.
ACCESSION   AA572148
VERSION     AA572148.1  GI:2345343
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:556422
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 279.
FEATURES             Location/Qualifiers
     source          1. .313
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:975694"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; 5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' 3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      113 a     66 c     45 g     89 t
ORIGIN

Alignment Scores:
Pred. No.:              670            Length:        313
Score:                  37.00          Matches:       8
Percent Similarity:     90.00%         Conservative:  1
Best Local Similarity:  80.00%         Mismatches:    1
Query Match:            75.51%         Indels:        0
DB:                     9              Gaps:          0

US-09-823-649A-6 (1-11) x AA572148 (1-313)

Qy      1 LeuSerLysArgIleGlyLeuSerValSer 10
          |||   :::|||||||||||||||||||||
Db    184 TTAGGCAGAAGGATAGGCTTGAGTGTCTCT 155

RESULT 15
AU112697/c
LOCUS       AU112697                350 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  AU112697 unpublished oligo-capped cDNA library Caenorhabditis
            elegans cDNA clone yk745h9 5', mRNA sequence.
ACCESSION   AU112697
VERSION     AU112697.1  GI:10926264
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 350)
  AUTHORS   Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
            and Sugano,S.
  TITLE     A complementary view of the C. elegans genome
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yuji Kohara
            Genome Biology Lab.
            National Institute of Genetics
            Yata 1111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykohara@lab.nig.ac.jp.
FEATURES             Location/Qualifiers
```

```
     source          1. .350
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /clone="yk745h9"
                     /clone_lib="unpublished oligo-capped cDNA library"
                     /sex="Hermaphrodite"
                     /tissue_type="whole animal"
                     /dev_stage="varied"
                     /note="Organ: ovary; Vector: Lambd Zap II; Approximately
                     1,000 ovaries were hand dissected from adult C. elegans.
                     RNA was extracted and the cDNA was amplified using the
                     SMART PCR cDNA Library Construction Kit (Clontech). This
                     was ligated into Lambda Zap."
BASE COUNT       95 a     79 c     86 g     90 t
ORIGIN

Alignment Scores:
Pred. No.:              778           Length:          350
Score:                  37.00         Matches:         7
Percent Similarity:     100.00%       Conservative:    4
Best Local Similarity:  63.64%        Mismatches:      0
Query Match:            75.51%        Indels:          0
DB:                     9             Gaps:            0

US-09-823-649A-6 (1-11) x AU112697 (1-350)

Qy       1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
           ::::::|||:::||||||||||||||||||:::
Db     212 ATCAACAAGAAGATTGGCTTATCAGTATCCAAA 180


Search completed: January 21, 2003, 12:24:26
Job time : 1172.14 secs
```

US-09-823-649A-5 (1-11) x US-09-815-242-9415 (1-2031)

```
Qy      1 LeuSerValArgLeuGly***ProValLysGlu 11
          ||||||||||::::::     |||:::||||||
Db     70 TTATCTGTTAAGGTGAGCTTACCTATAAAGGAA 102
```

RESULT 13
US-09-286-240-5/c
; Sequence 5, Application US/09286240
; Patent No. US20020010320A1
; GENERAL INFORMATION:
; APPLICANT: Fett, James W
; TITLE OF INVENTION: Chimeric and Humanized Antibodies to Angiogenin
; FILE REFERENCE: 10498/74073
; CURRENT APPLICATION NUMBER: US/09/286,240
; CURRENT FILING DATE: 1999-04-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1809)..(2252)
US-09-286-240-5

Alignment Scores:			
Pred. No.:	544	Length:	4668
Score:	34.00	Matches:	7
Percent Similarity:	87.50%	Conservative:	0
Best Local Similarity:	87.50%	Mismatches:	1
Query Match:	69.39%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-5 (1-11) x US-09-286-240-5 (1-4668)

```
Qy      1 LeuSerValArgLeuGly***Pro 8
          ||||||||||||||||||   |||
Db   4038 CTTTCAGTGAGGCTGGGCTCTCCT 4015
```

RESULT 14
US-09-863-777-1/c
; Sequence 1, Application US/09863777
; Patent No. US20020019359A1
; GENERAL INFORMATION:
; APPLICANT: Fett, James W.
; APPLICANT: Olson, Karen A.
; TITLE OF INVENTION: Antisense Inhibition of Angiogenin Expression
; FILE REFERENCE: 10498/05286
; CURRENT APPLICATION NUMBER: US/09/863,777
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/041182
; PRIOR FILING DATE: 1997-03-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1809)..(2252)
US-09-863-777-1

Alignment Scores:			
Pred. No.:	544	Length:	4668
Score:	34.00	Matches:	7
Percent Similarity:	87.50%	Conservative:	0
Best Local Similarity:	87.50%	Mismatches:	1
Query Match:	69.39%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-5 (1-11) x US-09-863-777-1 (1-4668)

```
Qy      1 LeuSerValArgLeuGly***Pro 8
          ||||||||||||||||||   |||
Db   4038 CTTTCAGTGAGGCTGGGCTCTCCT 4015
```

RESULT 15
US-09-880-107-2239/c
; Sequence 2239, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2239
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M11567
US-09-880-107-2239

Alignment Scores:			
Pred. No.:	544	Length:	4668
Score:	34.00	Matches:	7
Percent Similarity:	87.50%	Conservative:	0
Best Local Similarity:	87.50%	Mismatches:	1
Query Match:	69.39%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-5 (1-11) x US-09-880-107-2239 (1-4668)

```
Qy      1 LeuSerValArgLeuGly***Pro 8
          ||||||||||||||||||   |||
Db   4038 CTTTCAGTGAGGCTGGGCTCTCCT 4015
```

Search completed: January 21, 2003, 10:07:50
Job time : 33.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds
(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649A-5
Perfect score: 49
Sequence: 1 LSVRLGXPVKE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24616/app_query.fasta_1.1393
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0982364 _@CGN_1_1_2820_@runat_21012003_093151_24616 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOC -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		42	85.7	589	17	BH110033	BH110033 RPCI-24-3
2	c	41	83.7	584	14	BU008878	BU008878 QGH8O11.y
3		41	83.7	985	12	BE784762	BE784762 601473679
4		40	81.6	350	10	AW047855	AW047855 UI-M-BH1-
5		40	81.6	529	12	BG102214	BG102214 RHIZ2_22_
6		39	79.6	649	17	AZ795132	AZ795132 2M0049P12
7		39	79.6	658	12	BF214447	BF214447 601845635
8	c	39	79.6	720	17	BH699492	BH699492 BOMBK74TF
9	c	38	77.6	112	9	AA293830	AA293830 zt65b09.r
10	c	38	77.6	190	14	BQ565397	BQ565397 qi36q12.y
11	c	38	77.6	259	17	AZ769101	AZ769101 1M0569J24
12		38	77.6	300	13	BI690928	BI690928 603312076
13	c	38	77.6	309	17	BH743626	BH743626 gt25a07.b
14	c	38	77.6	338	10	AW477975	AW477975 17923 MAR
15		38	77.6	384	17	AZ465796	AZ465796 1M0275N19
16		38	77.6	405	17	CNS05R1K	AL349985 Tetraodon
17		38	77.6	415	10	AW772450	AW772450 hn74c08.x
18		38	77.6	432	9	AA537681	AA537681 vk44f02.r
19	c	38	77.6	517	10	AW785618	AW785618 l16849 MA
20	c	38	77.6	523	13	BI507078	BI507078 BB170023B
21		38	77.6	531	9	AI596792	AI596792 vk44f02.y
22		38	77.6	546	10	AW642017	AW642017 cml4a06.w
23	c	38	77.6	632	17	AZ460045	AZ460045 1M0265A18
24		38	77.6	638	13	BI186406	BI186406 UNL-P-FN-
25		38	77.6	674	17	AG064313	AG064313 Pan trog1
26		38	77.6	685	12	BF702429	BF702429 MI-P-A2-a
27		38	77.6	699	13	BI184838	BI184838 UNL-P-FN-
28	c	38	77.6	716	17	AG076995	AG076995 Pan trog1
29		37	75.5	310	9	AV154670	AV154670 AV154670
30		37	75.5	326	12	BG661863	BG661863 kx07c12.y
31	c	37	75.5	366	13	BI344292	BI344292 372728 MA
32		37	75.5	371	17	AZ319685	AZ319685 1M0039G21
33	c	37	75.5	417	17	AQ338684	AQ338684 HS_2210_B
34		37	75.5	450	10	AW387729	AW387729 MR4-ST011
35		37	75.5	468	12	BF387658	BF387658 UI-R-CA1-
36		37	75.5	475	9	AA564641	AA564641 nj25e12.s
37	c	37	75.5	496	10	AW581884	AW581884 MR4-ST011
38	c	37	75.5	530	9	AI562393	AI562393 TENS2383
39	c	37	75.5	533	14	BM884318	BM884318 rc21h03.y
40	c	37	75.5	690	12	BF033871	BF033871 601456259
41		37	75.5	691	17	BH722607	BH722607 BOMCL02TF
42	c	37	75.5	742	14	BQ573626	BQ573626 UI-M-FD0-
43	c	37	75.5	979	17	CNS02EIJ	AL193780 Tetraodon
44	c	37	75.5	1004	12	BG213266	BG213266 RST32872
45		37	75.5	1052	13	BM548113	BM548113 AGENCOURT

ALIGNMENTS

RESULT 1
BH110033
LOCUS BH110033 589 bp DNA linear GSS 19-JUL-2001
DEFINITION RPCI-24-323K23.TJ RPCI-24 Mus musculus genomic clone RPCI-24-323K23 , DNA sequence.
ACCESSION BH110033
VERSION BH110033.1 GI:14943718
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 589)
AUTHORS Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M., Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebreqeorgis,E.,

```
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-24-323K23.TVB
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 323  row: K  column: 23
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .589
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-323K23"
                     /clone_lib="RPCI-24"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was  cloned in the pTARBAC1 cloning vector at the
                     BamH1 sites using MboI  partially digested male C57BL/6J
                     DNA."
BASE COUNT      157 a    137 c    170 g    125 t
ORIGIN

Alignment Scores:
Pred. No.:                534           Length:          589
Score:                    42.00         Matches:         8
Percent Similarity:       81.82%        Conservative:    1
Best Local Similarity:    72.73%        Mismatches:      2
Query Match:              85.71%        Indels:          0
DB:                       17            Gaps:            0

US-09-823-649A-5 (1-11) x BH110033 (1-589)

Qy       1 LeuSerValArgLeuGly***ProValLysGlu 11
           |||||||||||||||||   |||:::   |||
Db      61 CTATCAGTGAGGCTGGGAAGCCCCATCAGTGAG 93

RESULT 2
BU008878/c
LOCUS       BU008878                584 bp    mRNA    linear   EST 22-AUG-2002
DEFINITION  QGH8011.yg.ab1 QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
            QGH8011, mRNA sequence.
ACCESSION   BU008878
VERSION     BU008878.1  GI:22443285
KEYWORDS    EST.
SOURCE      Lactuca sativa.
  ORGANISM  Lactuca sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QG_CA_Contig6932, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QGH8  row: O  column: 11.
FEATURES             Location/Qualifiers
     source          1. .584
                     /organism="Lactuca sativa"
                     /cultivar="L.serriola"
                     /db_xref="taxon:4236"
                     /clone="QGH8011"
                     /clone_lib="QG_EFGHJ lettuce serriola"
                     /lab_host="E.coli"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 10 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_SEQ-Not found"
BASE COUNT      170 a    123 c    131 g    160 t
ORIGIN

Alignment Scores:
Pred. No.:                804           Length:          584
Score:                    41.00         Matches:         9
Percent Similarity:       81.82%        Conservative:    0
Best Local Similarity:    81.82%        Mismatches:      2
Query Match:              83.67%        Indels:          0
DB:                       14            Gaps:            0

US-09-823-649A-5 (1-11) x BU008878 (1-584)

Qy       1 LeuSerValArgLeuGly***ProValLysGlu 11
           |||||||||||||   |||   ||||||||||||
Db     213 CTATCTGTAAGGCCAGGGGCACCGGTGAAAGAG 181

RESULT 3
BE784762
LOCUS       BE784762                985 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601473679F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876691 5',
            mRNA sequence.
ACCESSION   BE784762
VERSION     BE784762.1  GI:10205960
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 985)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9637  row: n  column: 20
            High quality sequence stop: 287.
FEATURES             Location/Qualifiers
     source          1. .985
```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3876691"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies. "

BASE COUNT 290 a 312 c 254 g 129 t
ORIGIN

Alignment Scores:

Pred. No.:	1.34e+03	Length:	985
Score:	41.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	83.67%	Indels:	0
DB:	12	Gaps:	0

US-09-823-649A-5 (1-11) x BE784762 (1-985)

```
Qy      1 LeuSerValArgLeuGly***ProValLysGlu 11
          ||||||:::|||:::|||   |||   ||||||
Db    283 CTCAGCATCCGTATTGGCACACCAACCAAAGAG 315
```

RESULT 4
AW047855
LOCUS AW047855 350 bp mRNA linear EST 18-SEP-1999
DEFINITION UI-M-BH1-als-e-04-0-UI.s1 NIH_BMAP_M_S2 Mus musculus cDNA clone UI-M-BH1-als-e-04-0-UI 3', mRNA sequence.
ACCESSION AW047855
VERSION AW047855.1 GI:5908384
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 350)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.
Seq primer: M13 Forward
POLYA=No.

FEATURES Location/Qualifiers
source 1. .350
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH1-als-e-04-0-UI"
/clone_lib="NIH_BMAP_M_S2"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NIH_BMAP_M_S2 library is a subtracted library derived from NIH_BMAP_M_S1, which in turn is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hipoccampus). The driver used for subtraction consisted of a pool of 5,000 clones from the NIH_BMAP_M_S1 library and a pool of 2,000 clones obtained from non-normalized and normalized mouse brain spinal cord libraries.
TAG_LIB=NIH_BMAP_M_S2
TAG_TISSUE=prefrontal-cortex
TAG_SEQ=GCTCA"

BASE COUNT 76 a 107 c 98 g 69 t
ORIGIN

Alignment Scores:

Pred. No.:	737	Length:	350
Score:	40.00	Matches:	8
Percent Similarity:	81.82%	Conservative:	1
Best Local Similarity:	72.73%	Mismatches:	2
Query Match:	81.63%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-5 (1-11) x AW047855 (1-350)

```
Qy      1 LeuSerValArgLeuGly***ProValLysGlu 11
            |||   ||||||:::|||   ||||||||||||
Db    234 CTACCTGTTCGAATGGGATACCCAGTTAAAGAA 266
```

RESULT 5
BG102214
LOCUS BG102214 529 bp mRNA linear EST 30-JAN-2001
DEFINITION RHIZ2_22_B12.b1_A003 Rhizome2 (RHIZ2) Sorghum propinquum cDNA, mRNA sequence.
ACCESSION BG102214
VERSION BG102214.1 GI:12617047
KEYWORDS EST.
SOURCE Sorghum propinquum.
ORGANISM Sorghum propinquum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 529)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Paterson,A., Sudman,M. and Pratt,L.H.
TITLE An EST database from Sorghum: Sorghum propinquum rhizomes
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0219
Email: mmpratt@uga.edu
Seq primer: JEN REV
High quality sequence stop: 499
POLYA=No.

FEATURES Location/Qualifiers
source 1. .529
/organism="Sorghum propinquum"
/db_xref="taxon:132711"
/clone_lib="Rhizome2 (RHIZ2)"
/note="Organ: Rhizomes; Vector: pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."

BASE COUNT 78 a 162 c 180 g 109 t
ORIGIN

Alignment Scores:

Pred. No.:	1.11e+03	Length:	529
Score:	40.00	Matches:	8
Percent Similarity:	81.82%	Conservative:	1
Best Local Similarity:	72.73%	Mismatches:	2

```
Query Match:            81.63%      Indels:        0
DB:                     12          Gaps:          0

US-09-823-649A-5 (1-11) x BG102214 (1-529)

Qy      1 LeuSerValArgLeuGly***ProValLysGlu 11
          |||   |||||||||||   ||||||:::|||
Db    444 TTGGACGTGAGATTAGGAGCTCCAGTGGAGGAG 476

RESULT 6
AZ795132
LOCUS       AZ795132                649 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0049P12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0049P12 F, DNA sequence.
ACCESSION   AZ795132
VERSION     AZ795132.1  GI:12941837
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 649)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0049  row: P  column: 12
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 649.
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0049P12"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      181 a    149 c    174 g    145 t
ORIGIN

Alignment Scores:
Pred. No.:              2.05e+03    Length:        649
Score:                  39.00       Matches:       8
Percent Similarity:     80.00%      Conservative:  0
Best Local Similarity:  80.00%      Mismatches:    2
Query Match:            79.59%      Indels:        0
DB:                     17          Gaps:          0

US-09-823-649A-5 (1-11) x AZ795132 (1-649)

Qy      2 SerValArgLeuGly***ProValLysGlu 11
          ||||||||||||||   ||||||   |||
Db     74 TCAGTGAGGCTGGGAAGCCCCGTCAGTGAG 103

RESULT 7
BF214447
LOCUS       BF214447                658 bp    mRNA    linear   EST 06-NOV-2000
DEFINITION  601845635F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4076553 5',
            mRNA sequence.
ACCESSION   BF214447
VERSION     BF214447.1  GI:11108033
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM930  row: f  column: 10
            High quality sequence stop: 416.
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4076553"
                     /clone_lib="NIH_MGC_55"
                     /tissue_type="from acute myelogenous leukemia"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); Double-stranded cDNA was prepared from cell line RNA.
                     5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.65 kb (range 0.9-4.0 kb).  14/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      142 a    166 c    212 g    138 t
ORIGIN

Alignment Scores:
Pred. No.:              2.08e+03    Length:        658
Score:                  39.00       Matches:       7
Percent Similarity:     90.91%      Conservative:  3
Best Local Similarity:  63.64%      Mismatches:    1
Query Match:            79.59%      Indels:        0
DB:                     12          Gaps:          0

US-09-823-649A-5 (1-11) x BF214447 (1-658)
```

```
Qy       1 LeuSerValArgLeuGly***ProValLysGlu 11
           :::|||:::|||||||||   |||:::||||||
Db     384 ATGAGCCTGAGATTGGGGAGGCCCCTGAAGGAG 416

RESULT 8
BH699492/c
LOCUS       BH699492                720 bp    DNA     linear   GSS 20-FEB-2002
DEFINITION  BOMBK74TF BO_2_3_KB Brassica oleracea genomic clone BOMBK74, DNA
            sequence.
ACCESSION   BH699492
VERSION     BH699492.1  GI:18774833
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 720)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOMBK74TR
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .720
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOMBK74"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      164 a    239 c    147 g    170 t
ORIGIN

Alignment Scores:
Pred. No.:                2.27e+03        Length:         720
Score:                    39.00           Matches:        8
Percent Similarity:       90.00%          Conservative:   1
Best Local Similarity:    80.00%          Mismatches:     1
Query Match:              79.59%          Indels:         0
DB:                       17              Gaps:           0

US-09-823-649A-5 (1-11) x BH699492 (1-720)

Qy       1 LeuSerValArgLeuGly***ProValLys 10
           :::|||||||||||||||   |||||||||
Db     135 GTGAGTGTTAGACTCGGTTTCCCGGTGAAG 106

RESULT 9
AA293830/c
LOCUS       AA293830                112 bp    mRNA    linear   EST 12-AUG-1997
DEFINITION  zt65b09.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:727193
            5' similar to SW:MLH1_HUMAN P40692 MUTL PROTEIN HOMOLOG 1 ;, mRNA
            sequence.
ACCESSION   AA293830
VERSION     AA293830.1  GI:1941753
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 112)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 645   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .112
                     /organism="Homo sapiens"
                     /db_xref="GDB:5924105"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:727193"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with  a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       28 a     24 c     29 g     31 t
ORIGIN

Alignment Scores:
Pred. No.:                553             Length:         112
Score:                    38.00           Matches:        6
Percent Similarity:       90.00%          Conservative:   3
Best Local Similarity:    60.00%          Mismatches:     1
Query Match:              77.55%          Indels:         0
DB:                       9               Gaps:           0

US-09-823-649A-5 (1-11) x AA293830 (1-112)

Qy       1 LeuSerValArgLeuGly***ProValLys 10
           |||||||||||||||||   |||:::|||
Db      39 TTGTCAATCAGAATGGGTAATCCAATCAAG 10

RESULT 10
BQ565397/c
LOCUS       BQ565397                190 bp    mRNA    linear   EST 19-JUN-2002
DEFINITION  gi36g12.y1 Mouse Organ of Corti cDNA pBluescript Mus musculus cDNA
            clone gi36g12 5', mRNA sequence.
ACCESSION   BQ565397
VERSION     BQ565397.1  GI:21468714
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 190)
  AUTHORS   Kachar,B.
  TITLE     EST analysis of gene expression in the mouse Organ of Corti at the
            onset of hearing
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kachar,B.
            Structural Cell Biology
            National Institute of Deafness and other Communication Disorders
```

50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
Tel: 301-402-1599
Fax: 301-402-1765
Email: kacharb@nidcd.nih.gov
Plate: 36 row: g column: 12
Seq primer: M13RP1 reverse primer (ABI).

FEATURES Location/Qualifiers
source 1. .190
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="gi36g12"
/clone_lib="Mouse Organ of Corti cDNA pBluescript"
/sex="male and female"
/dev_stage="Post natal day 5 to 13"
/note="Organ: Organ of Corti; Vector: pBluescript; The organ of Corti (OC) was fine dissected from a total of 386 OC as follows: 102 samples from post-natal (P) day 5; 72 from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10; 14 from P12 and 24 from P13. After killing animals by cervical dislocation followed by decapitation, the bulla was removed and opened in Leibowitz medium. The bony capsule of the cochlea was chipped away, stria vascularis and spiral ligament were removed and the sensory epithelium was carefully dissected out of the modiolus. Total RNA was extracted using the micro Fasttrack kit (catalog # K1593-02; Invitrogen, Carlsbad, CA), according to manufacturer's instructions. Reverse transcription and library construction were carried out with the Uni-Zap XR vector kit (catalog # 237211, Stratagene) and Uni-Zap XR Gigapack III Gold Cloning kit (catalog # 237612), both from Stratagene (La Jolla, CA, USA), according to manufacturer's instructions. Briefly: 1.5 ug mRNA was reverse transcribed using a hybrid oligo(dT) linker-primer that contains an Xho I site. First strand synthesis was primed with the linker- primer and transcribed using Moloney murine leukemia virus reverse transcriptase (MMLV-RT) and 5-methyl dCTP. The second strand was synthesized with DNA polymerase and RNase H. Complementary DNA was blunt ended with Pfu DNA polymerase, ligated with EcoR I adapters in the presence of ligase and digested with Xho I. The cDNA was sequentially size fractionated over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden) and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA) columns to enrich for cDNAs greater than 400bp and 1000 bp , respectively. The cDNA was then directionally ligated to the Uni-ZAP XR vector, which had been predigested with EcoR I and Xho I. The phagemid was packaged with Gigapak III Gold and, upon titration on XL1 Blue MRF' cells, the yield of the phage library was estimated to be 11,100,000 recombinants. Stratagene's ExAssist interference resistance helper phage (catalogue # 211203) was adopted to rescue plasmid DNA from the phages. Upon plating of the rescued library, individual cDNA clones were selected and grown in 96-well, 2 ml growth plate. Plasmid DNA was purified from 200 ul of saturated culture with the Concert96(TM) plasmid purification kit (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. ESTs from the 5' end of the cDNA clones were generated with the universal M13 reverse primer (CAGGAAACAGCTATGACC) and 25% strength BigDye terminator sequencing chemistry (Applied Biosystems, Foster City, CA). Sequencing reactions were performed on MJ Tetrad thermal cyclers (MJ Research, Waltham, MA), and analyzed on 3700 automated capillary sequencers using POP5 polymer (Applied Biosystems, Foster City, CA). The frequency distribution of the library is as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10; 1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of genes are present in GenBank and have know function; 23% have hits in GenBank, but do not have assigned function; 12% are uncharacterized ESTs and 20% are unidentified."

BASE COUNT 45 a 57 c 48 g 40 t
ORIGIN

Alignment Scores:

Pred. No.:	930	Length:	190
Score:	38.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	77.55%	Indels:	0
DB:	14	Gaps:	0

US-09-823-649A-5 (1-11) x BQ565397 (1-190)

```
Qy      1 LeuSerValArgLeuGly***ProValLysGlu 11
          ||||||:::||||||      |||||||||:::
Db    146 TTGAGCATCCGTTTGTCTGCTCCTGTCAAGCAG 114
```

RESULT 11
AZ769101/c
LOCUS AZ769101 259 bp DNA linear GSS 16-FEB-2001
DEFINITION 1M0569J24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0569J24 F, DNA sequence.
ACCESSION AZ769101
VERSION AZ769101.1 GI:12888888
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 259)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0569 row: J column: 24
Seq primer: CGTTGTAAAACCACGGCCAGT
Class: plasmid ends
High quality sequence stop: 259.

FEATURES Location/Qualifiers
source 1. .259
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0569J24"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to

adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 60 a 55 c 60 g 84 t
ORIGIN

Alignment Scores:

Pred. No.:	1.26e+03	Length:	259
Score:	38.00	Matches:	8
Percent Similarity:	81.82%	Conservative:	1
Best Local Similarity:	72.73%	Mismatches:	2
Query Match:	77.55%	Indels:	0
DB:	17	Gaps:	0

US-09-823-649A-5 (1-11) x AZ769101 (1-259)

```
Qy      1 LeuSerValArgLeuGly***ProValLysGlu 11
          |||   ||||||||||||   |||||||||:::
Db    194 CTCTGGGTAAGACTGGGAAGCCCAGTGAAGGAC 162
```

RESULT 12
BI690928
LOCUS BI690928 300 bp mRNA linear EST 18-SEP-2001
DEFINITION 603312076F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5352234 5', mRNA sequence.
ACCESSION BI690928
VERSION BI690928.1 GI:15653557
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 300)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11895 row: g column: 19
High quality sequence stop: 300.

```
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5352234"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT       60 a     91 c     75 g     74 t
ORIGIN
```

Alignment Scores:

Pred. No.:	1.46e+03	Length:	300
Score:	38.00	Matches:	6
Percent Similarity:	90.91%	Conservative:	4
Best Local Similarity:	54.55%	Mismatches:	1
Query Match:	77.55%	Indels:	0
DB:	13	Gaps:	0

US-09-823-649A-5 (1-11) x BI690928 (1-300)

```
Qy      1 LeuSerValArgLeuGly***ProValLysGlu 11
          ||||||:::|||||||||   |||:::::::::
Db     41 CTTTCCATCAGGCTAGGGAATCCTATCAGGAAG 73
```

RESULT 13
BH743626/c
LOCUS BH743626 309 bp DNA linear GSS 25-FEB-2002
DEFINITION gt25a07.b1 BoBuds01 Brassica oleracea genomic clone gt25a07 5', DNA sequence.
ACCESSION BH743626
VERSION BH743626.1 GI:18878240
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 309)
AUTHORS Katari,M., O'Shaughnessy,A., Palmer,L., Bahret,A., Baker,J., Balija,V., Cunnius,D.M., Katzenberger,F., King,L., Kirchoff,K., Kuit,K., Miller,B., Muller,S., Nascimento,L., Preston,R., Santos,L., Shah,R., Zutavern,T., Dedhia,N., Rabinowicz,P.D. and McCombie,W.R.
TITLE Whole Genome Shotgun Reads from Brassica oleracia (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mccombie@cshl.org
Plate: gt25 row: a column: 07
Seq primer: -21UnivFwd
Class: shotgun
High quality sequence stop: 309.

```
FEATURES             Location/Qualifiers
     source          1. .309
                     /organism="Brassica oleracea"
                     /db_xref="taxon:3712"
                     /clone="gt25a07"
                     /clone_lib="BoBuds01"
                     /note="Vector: M13 for .x reads, pBluescript for .b and .g
                     reads; Site_1: EcoRV; Whole genome shotgun library from
                     flowering buds. DNA was purified from a crude nuclear prep
                     using Brassica oleracea TO1000DH3 buds provided by Thomas
                     Osborn at the University of Wisconsin. Genomic DNA
                     provided by Pablo Rabinowicz (CSHL) and shotgun library
                     prepared in McCombie Lab."
BASE COUNT       64 a     78 c     71 g     96 t
ORIGIN
```

Alignment Scores:

Pred. No.:	1.5e+03	Length:	309
Score:	38.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	77.55%	Indels:	0
DB:	17	Gaps:	0

US-09-823-649A-5 (1-11) x BH743626 (1-309)

```
Qy      1 LeuSerValArgLeuGly***ProValLysGlu 11
          :::|||:::||||||      ||||||||||||
Db    114 ATATCGATCCGACTTTCAGAACCCGTCAAAGAA 82
```

RESULT 14
AW477975/c
LOCUS AW477975 338 bp mRNA linear EST 09-JUL-2000
DEFINITION 17923 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION AW477975
VERSION AW477975.1 GI:7048081

KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 338)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 8 row: P column: 8
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .338
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."
BASE COUNT 87 a 84 c 95 g 72 t
ORIGIN

Alignment Scores:

Pred. No.:	1.64e+03	Length:	338
Score:	38.00	Matches:	8
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	1
Query Match:	77.55%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-5 (1-11) x AW477975 (1-338)

```
Qy       1 LeuSerValArgLeuGly***ProVal 9
           ||||||||||||||||||   ||||||
Db     323 CTGTCAGTCAGGTTGGGCTCTCCTGTG 297
```

RESULT 15
AZ465796
LOCUS AZ465796 384 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0275N19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0275N19 R, DNA sequence.
ACCESSION AZ465796
VERSION AZ465796.1 GI:10623921
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 384)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0275 row: N column: 19
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 384.
FEATURES Location/Qualifiers
source 1. .384
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0275N19"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 108 a 96 c 89 g 91 t
ORIGIN

Alignment Scores:

Pred. No.:	1.86e+03	Length:	384
Score:	38.00	Matches:	8
Percent Similarity:	81.82%	Conservative:	1
Best Local Similarity:	72.73%	Mismatches:	2
Query Match:	77.55%	Indels:	0
DB:	17	Gaps:	0

US-09-823-649A-5 (1-11) x AZ465796 (1-384)

```
Qy       1 LeuSerValArgLeuGly***ProValLysGlu 11
           |||  ||||||||||||  |||||||||:::
Db     310 CTCTGGGTAAGACTGGGAAGCCCAGTGAAGGAC 342
```

Search completed: January 21, 2003, 12:24:22
Job time : 1172.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds
(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649A-1
Perfect score: 18
Sequence: 1 LXXXXXXXXXE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24626/app_query.fasta_1.1393
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649_@CGN_1_1_77_@runat_21012003_093151_24626
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	9	50.0	257	10	US-09-878-574-15531	Sequence 15531, A
	2	9	50.0	415	10	US-09-878-574-4032	Sequence 4032, Ap
c	3	9	50.0	462	10	US-09-864-761-95	Sequence 95, Appl
c	4	9	50.0	618	10	US-09-864-761-16934	Sequence 16934, A
	5	9	50.0	618	10	US-09-864-761-18952	Sequence 18952, A
c	6	9	50.0	683	9	US-10-040-739-303	Sequence 303, App
	7	9	50.0	989	10	US-09-864-761-2212	Sequence 2212, Ap
	8	9	50.0	1556	9	US-09-981-353-64	Sequence 64, Appl
c	9	9	50.0	1842	9	US-09-981-876-113	Sequence 113, App
c	10	9	50.0	1903	9	US-09-981-876-55	Sequence 55, Appl
	11	9	50.0	2037	10	US-09-822-830A-25	Sequence 25, Appl
c	12	9	50.0	2281	10	US-09-799-777-101	Sequence 101, App
	13	9	50.0	2793	10	US-09-880-107-3437	Sequence 3437, Ap
c	14	9	50.0	3501	9	US-09-938-842A-13	Sequence 13, Appl
	15	9	50.0	3531	9	US-09-992-598-32	Sequence 32, Appl
	16	9	50.0	3531	9	US-09-989-293A-32	Sequence 32, Appl
	17	9	50.0	3531	9	US-09-989-735-32	Sequence 32, Appl
	18	9	50.0	3531	9	US-09-990-444-32	Sequence 32, Appl
	19	9	50.0	3531	9	US-09-989-730-32	Sequence 32, Appl
	20	9	50.0	3531	9	US-09-990-436-32	Sequence 32, Appl
	21	9	50.0	3531	9	US-09-991-181-32	Sequence 32, Appl
	22	9	50.0	3531	9	US-09-993-687-32	Sequence 32, Appl
	23	9	50.0	3531	9	US-09-989-734-32	Sequence 32, Appl
	24	9	50.0	3531	9	US-09-997-653-32	Sequence 32, Appl
	25	9	50.0	3531	10	US-09-989-722-32	Sequence 32, Appl
	26	9	50.0	3531	10	US-09-989-723-32	Sequence 32, Appl
	27	9	50.0	3531	10	US-09-989-279-32	Sequence 32, Appl
	28	9	50.0	3531	10	US-09-989-727-32	Sequence 32, Appl
	29	9	50.0	3531	10	US-09-989-731-32	Sequence 32, Appl
	30	9	50.0	3531	10	US-09-989-732-32	Sequence 32, Appl
	31	9	50.0	3531	10	US-09-991-073-32	Sequence 32, Appl
	32	9	50.0	3531	10	US-09-990-442-32	Sequence 32, Appl
	33	9	50.0	3531	10	US-09-991-163-32	Sequence 32, Appl
	34	9	50.0	3531	10	US-09-993-604-32	Sequence 32, Appl
	35	9	50.0	3531	10	US-09-990-456-32	Sequence 32, Appl
	36	9	50.0	3531	10	US-09-989-721-32	Sequence 32, Appl
	37	9	50.0	4100	9	US-09-813-718-5	Sequence 5, Appli
c	38	9	50.0	4273	9	US-09-974-298-143	Sequence 143, App
	39	9	50.0	5174	9	US-09-813-718-1	Sequence 1, Appli
	40	9	50.0	13444	10	US-09-764-877-3660	Sequence 3660, Ap
	41	9	50.0	17967	10	US-09-764-848-50	Sequence 50, Appl
	42	9	50.0	18871	10	US-09-764-847-1317	Sequence 1317, Ap
c	43	8	44.4	82	10	US-09-878-574-7261	Sequence 7261, Ap
c	44	8	44.4	97	10	US-09-864-761-20170	Sequence 20170, A
c	45	8	44.4	194	9	US-09-900-379-107	Sequence 107, App

ALIGNMENTS

RESULT 1
US-09-878-574-15531
; Sequence 15531, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 15531
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701070216H1
US-09-878-574-15531

Alignment Scores:

Pred. No.:	2.91	Length:	257
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9

```
Query Match:            50.00%      Indels:        0
DB:                     10          Gaps:          0

US-09-823-649A-1 (1-11) x US-09-878-574-15531 (1-257)

Qy      1 Leu**************************Glu 11
          |||                          |||
Db      8 CTGACCTCAGCGGCGGCGGCGGCGGCAACAGAA 40

RESULT 2
US-09-878-574-4032
; Sequence 4032, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFFRENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 4032
;   LENGTH: 415
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: LIB3028-002-Q1-B1-E10
US-09-878-574-4032

Alignment Scores:
Pred. No.:              3.59        Length:        415
Score:                  9.00        Matches:       2
Percent Similarity:     18.18%      Conservative:  0
Best Local Similarity:  18.18%      Mismatches:    9
Query Match:            50.00%      Indels:        0
DB:                     10          Gaps:          0

US-09-823-649A-1 (1-11) x US-09-878-574-4032 (1-415)

Qy      1 Leu**************************Glu 11
          |||                          |||
Db     13 CTGACCTCAGCGGCGGCGGCGGCGGCAACAGAA 45

RESULT 3
US-09-864-761-95/c
; Sequence 95, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 95
;   LENGTH: 462
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007051.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.82
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
US-09-864-761-95

Alignment Scores:
Pred. No.:              3.77        Length:        462
Score:                  9.00        Matches:       2
Percent Similarity:     18.18%      Conservative:  0
Best Local Similarity:  18.18%      Mismatches:    9
Query Match:            50.00%      Indels:        0
DB:                     10          Gaps:          0

US-09-823-649A-1 (1-11) x US-09-864 761-95 (1-462)

Qy      1 Leu**************************Glu 11
          |||                          |||
Db    288 CTGACAGCAGCAGCCTCTTCCACATCAAGTGAA 256

RESULT 4
US-09-864-761-16934/c
; Sequence 16934, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
```

```
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16934
;   LENGTH: 618
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007051.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.82
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
;   OTHER INFORMATION: NT HIT: AB037784.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P22760, EVALUE 2.00e-42
;   OTHER INFORMATION: EST_HUMAN HIT: BE154994.1, EVALUE 0.00e+00
US-09-864-761-16934

Alignment Scores:
Pred. No.:              4.28          Length:          618
Score:                  9.00          Matches:         2
Percent Similarity:     18.18%        Conservative:    0
Best Local Similarity:  18.18%        Mismatches:      9
Query Match:            50.00%        Indels:          0
DB:                     10            Gaps:            0

US-09-823-649A-1 (1-11) x US-09-864-761-16934 (1-618)

Qy       1 Leu**************************Glu 11
             |||                          |||
Db     239 CTGACAGCAGCAGCCTCTTCCACATCAAGTGAA 207

RESULT 5
US-09-864-761-18952
```

```
;  Sequence 18952, Application US/09864761
;  Patent No. US20020048763A1
;  GENERAL INFORMATION:
;   APPLICANT: Penn, Sharron G.
;   APPLICANT:  Rank, David R.
;   APPLICANT:  Hanzel, David K.
;   APPLICANT:  Chen, Wensheng
;   TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;   TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;   FILE REFERENCE: Acomica-X-1
;   CURRENT APPLICATION NUMBER: US/09/864,761
;   CURRENT FILING DATE:  2001-05-23
;   PRIOR APPLICATION NUMBER: US 60/180,312
;   PRIOR FILING DATE: 2000-02-04
;   PRIOR APPLICATION NUMBER: US 60/207,456
;   PRIOR FILING DATE: 2000-05-26
;   PRIOR APPLICATION NUMBER: US 09/632,366
;   PRIOR FILING DATE: 2000-08-03
;   PRIOR APPLICATION NUMBER: GB 24263.6
;   PRIOR FILING DATE: 2000-10-04
;   PRIOR APPLICATION NUMBER: US 60/236,359
;   PRIOR FILING DATE: 2000-09-27
;   PRIOR APPLICATION NUMBER: PCT/US01/00666
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00667
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00664
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00669
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00665
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00668
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00663
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00662
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00661
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: PCT/US01/00670
;   PRIOR FILING DATE: 2001-01-30
;   PRIOR APPLICATION NUMBER: US 60/234,687
;   PRIOR FILING DATE: 2000-09-21
;   PRIOR APPLICATION NUMBER: US 09/608,408
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: US 09/774,203
;   PRIOR FILING DATE: 2001-01-29
;   NUMBER OF SEQ ID NOS: 49117
;   SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18952
;   LENGTH: 618
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007919.18
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: NT HIT: AB037784.1, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BE154994.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P22760, EVALUE 6.00e-39
US-09-864-761-18952

Alignment Scores:
Pred. No.:              4.28          Length:          618
```

```
Score:                 9.00          Matches:       2
Percent Similarity:    18.18%        Conservative:  0
Best Local Similarity: 18.18%        Mismatches:    9
Query Match:           50.00%        Indels:        0
DB:                    10            Gaps:          0

US-09-823-649A-1 (1-11) x US-09-864-761-18952 (1-618)

Qy      1 Leu**************************Glu 11
          |||                           |||
Db    324 CTGACAGCAGCAGCCTCTTCCACATCAAGTGAA 356

RESULT 6
US-10-040-739-303/c
; Sequence 303, Application US/10040739
; Patent No. US20020173635A1
;   GENERAL INFORMATION:
;        APPLICANT: Jacobs, Kenneth
;                   McCoy, John
;                   LaVallie, Edward
;                   Racie, Lisa
;                   Merberg, David
;                   Treacy, Maurice
;                   Spaulding, Vikki
;        TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;        NUMBER OF SEQUENCES: 1519
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy Disk
;             COMPUTER: IBM PC Compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/040,739
;             FILING DATE: 07-Jan-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/036,520
;             FILING DATE: 03-JUN-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brown, Scott A.
;             REGISTRATION NUMBER: 32,724
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 498-8224
;             TELEFAX: (617) 876-5851
;   INFORMATION FOR SEQ ID NO: 303:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 683 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 303:
US-10-040-739-303

Alignment Scores:
Pred. No.:             4.47          Length:        683
Score:                 9.00          Matches:       2
Percent Similarity:    18.18%        Conservative:  0
Best Local Similarity: 18.18%        Mismatches:    9
Query Match:           50.00%        Indels:        0
DB:                    9             Gaps:          0

US-09-823-649A-1 (1-11) x US-10-040-739-303 (1-683)

Qy      1 Leu**************************Glu 11
          |||                           |||
Db    451 TTAGCTGCCACATCCTCTTCCTCTTCCTCCGAG 419

RESULT 7
US-09-864-761-2212
; Sequence 2212, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2212
;   LENGTH: 989
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007919.18
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.6
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.9
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
US-09-864-761-2212

Alignment Scores:
```

```
Pred. No.:              5.27          Length:         989
Score:                  9.00          Matches:        2
Percent Similarity:     18.18%        Conservative:   0
Best Local Similarity:  18.18%        Mismatches:     9
Query Match:            50.00%        Indels:         0
DB:                     10            Gaps:           0

US-09-823-649A-1 (1-11) x US-09-864-761-2212 (1-989)

Qy      1 Leu**************************Glu 11
          |||                           |||
Db    676 CTGACAGCAGCAGCCTCTTCCACATCAAGTGAA 708

RESULT 8
US-09-981-353-64
; Sequence 64, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy W.
;  APPLICANT:  Jones, David A.
;  TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
;  FILE REFERENCE: PA-0038 US
;  CURRENT APPLICATION NUMBER: US/09/981,353
;  CURRENT FILING DATE:  2001-10-11
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 64
;   LENGTH: 1556
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020160382A1 3231154CB1
US-09-981-353-64

Alignment Scores:
Pred. No.:              6.43          Length:         1556
Score:                  9.00          Matches:        2
Percent Similarity:     18.18%        Conservative:   0
Best Local Similarity:  18.18%        Mismatches:     9
Query Match:            50.00%        Indels:         0
DB:                     9             Gaps:           0

US-09-823-649A-1 (1-11) x US-09-981-353-64 (1-1556)

Qy      1 Leu**************************Glu 11
          |||                           |||
Db   1054 TTGGCATCATCATCTACTGCTGCTGCTGCCGAG 1086

RESULT 9
US-09-981-876-113/c
; Sequence 113, Application US/09981876
; Patent No. US20020164669A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 70 Human Secreted Proteins
;  FILE REFERENCE: PZ001P1
;  CURRENT APPLICATION NUMBER: US/09/981,876
;  CURRENT FILING DATE:  2001-10-19
;  PRIOR APPLICATION NUMBER: 09/148,545
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: 60/040,162
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,333
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/038,621
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,161
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,626
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,334
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,336
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/040,163
;  PRIOR FILING DATE: 1997-03-07
;  PRIOR APPLICATION NUMBER: 60/047,615
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,600
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,597
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,502
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,633
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,583
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,617
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,618
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,503
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,592
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,581
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,584
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,500
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,587
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,492
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,598
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,613
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,582
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,596
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,612
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,632
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/047,601
;  PRIOR FILING DATE: 1997-05-23
;  PRIOR APPLICATION NUMBER: 60/043,580
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,568
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,314
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,569
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,311
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,671
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,674
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,669
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,312
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,313
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,672
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/043,315
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/048,974
```

; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/056,886
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,877
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,889
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,893
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,630
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,878
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,662
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,872
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,882
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,637
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,903
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,888
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,879
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,880
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,894
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,911
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,636
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,874
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,910
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,864
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,631
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,892
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/047,595
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/057,761
; PRIOR FILING DATE: 05-Sep-1997
; PRIOR APPLICATION NUMBER: 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,588
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,585
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,586
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,590
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,594
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,589
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,593
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,614
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,578
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,576
; PRIOR FILING DATE: 1997-04-11

; PRIOR APPLICATION NUMBER: 60/047,501
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,670
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/056,632
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,876
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,881
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 113
; LENGTH: 1842

Alignment Scores:

Pred. No.:	6.93	Length:	1842
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	9	Gaps:	0

US-09-823-649A-1 (1-11) x US-09-981-876-113 (1-1842)

```
Qy       1 Leu**************************Glu 11
             |||                          |||
Db     164 CTAACAACAGCKGCAGCAGCAGCAGCAGCAGAA 132
```

RESULT 10
US-09-981-876-55/c
; Sequence 55, Application US/09981876
; Patent No. US20020164669A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/981,876
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,161
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07

; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,581
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,584
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,500
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,613
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,582
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,596
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,612
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,568
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,569
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,311
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,671
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,674
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,669
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,312
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,313
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,672
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,315
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/056,886
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,877
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,889
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,893
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,630
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,878
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,662
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,872
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,882
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,637
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,903
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,888
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,879
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,880
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,894
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,911
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,636
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,874
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,910
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,864
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,631
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,892
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/047,595
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/057,791
; PRIOR FILING DATE: 05-Sep-1997
; PRIOR APPLICATION NUMBER: 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,588
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,585
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,586
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,590
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,594
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,589
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,593
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,614
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,578
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/047,501
; PRIOR FILING DATE: 1997-05-23

```
;  PRIOR APPLICATION NUMBER: 60/043,670
;  PRIOR FILING DATE: 1997-04-11
;  PRIOR APPLICATION NUMBER: 60/056,632
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: 60/056,664
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: 60/056,876
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: 60/056,881
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: 60/056,909
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: 60/056,875
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: 60/056,862
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: 60/056,887
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: 60/056,908
;  PRIOR FILING DATE: 1997-08-22
;  PRIOR APPLICATION NUMBER: 60/048,964
;  PRIOR FILING DATE: 1997-06-06
;  PRIOR APPLICATION NUMBER: 60/057,650
;  PRIOR FILING DATE: 1997-09-05
;  PRIOR APPLICATION NUMBER: 60/056,884
;  PRIOR FILING DATE: 1997-08-22
;  NUMBER OF SEQ ID NOS: 280
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
;   LENGTH: 1903

Alignment Scores:
Pred. No.:              7.03 <          Length:         1903
Score:                  9.00            Matches:        2
Percent Similarity:     18.18%          Conservative:   0
Best Local Similarity:  18.18%          Mismatches:     9
Query Match:            50.00%          Indels:         0
DB:                     9               Gaps:           0

US-09-823-649A-1 (1-11) x US-09-981-876-55 (1-1903)

Qy      1 Leu**************************Glu 11
            |||                          |||
Db    203 CTAACAACAGCGGCAGCAGCAGCAGCAGCAGAA 171

RESULT 11
US-09-822-830A-25
; Sequence 25  Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
;  APPLICANT: Genetics Institute, Inc.
;  APPLICANT:  Wong, Gordon G.
;  APPLICANT:  Clark, Hilary
;  APPLICANT:  Fechtel, Kim
;  APPLICANT:  Agostino, Michael J.
;  APPLICANT:  Howes, Steven H.
;  APPLICANT:  Resnick, Richard J.
;  APPLICANT:  Gulukota, Kamalakar
;  APPLICANT:  Graham, James R.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
;  FILE REFERENCE: GIN 6402
;  CURRENT APPLICATION NUMBER: US/09/822,830A
;  CURRENT FILING DATE:  2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195,604
;  PRIOR FILING DATE: 2000-04-06
;  NUMBER OF SEQ ID NOS: 631
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 2037
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-822-830A-25

Alignment Scores:
Pred. No.:              7.24            Length:         2037
Score:                  9.00            Matches:        2
Percent Similarity:     18.18%          Conservative:   0
Best Local Similarity:  18.18%          Mismatches:     9
Query Match:            50.00%          Indels:         0
DB:                     10              Gaps:           0

US-09-823-649A-1 (1-11) x US-09-822-830A-25 (1-2037)

Qy      1 Leu**************************Glu 11
            |||                          |||
Db     95 TTCGCCGCCGGCGCCGGCCGCCGGCAGCCGCCGCAG 127

RESULT 12
US-09-799-777-101/c
; Sequence 101, Application US/09799777
; Patent No. US20020091244A1
;   GENERAL INFORMATION:
;        APPLICANT: Lal, Preeti
;                   Hillman, Jennifer L.
;                   Corley, Neil C.
;                   Guegler, Karl J.
;                   Baugh, Mariah
;                   Sather, Susan
;                   Shah, Purvi
;        TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
;        NUMBER OF SEQUENCES: 154
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;             STREET: 3174 PORTER DRIVE
;             CITY: PALO ALTO
;             STATE: CALIFORNIA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/799,777
;             FILING DATE: 06-Mar-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/002,485
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BILLINGS, LUCY J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0459 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (650) 855-0555
;             TELEFAX: (650) 845-4166
;   INFORMATION FOR SEQ ID NO: 101:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2281 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: BRAITUT12
;             CLONE: 1615809
;        SEQUENCE DESCRIPTION: SEQ ID NO: 101 :
US-09-799-777-101

Alignment Scores:
Pred. No.:              7.61            Length:         2281
Score:                  9.00            Matches:        2
Percent Similarity:     18.18%          Conservative:   0
Best Local Similarity:  18.18%          Mismatches:     9
Query Match:            50.00%          Indels:         0
DB:                     10              Gaps:           0
```

US-09-823-649A-1 (1-11) x US-09-799-777-101 (1-2281)

```
Qy      1 Leu**************************Glu 11
          |||                           |||
Db    695 TTAGCTGCCACATCCTCTTCCTCTTCCTCCGAG 663
```

RESULT 13
US-09-880-107-3437
; Sequence 3437, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3437
; LENGTH: 2793
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U79725
US-09-880-107-3437

Alignment Scores:

Pred. No.:	8.32	Length:	2793
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-1 (1-11) x US-09-880-107-3437 (1-2793)

```
Qy      1 Leu**************************Glu 11
          |||                           |||
Db   1099 TTGGCATCATCATCTACTGCTGCTGCTGCCGAG 1131
```

RESULT 14
US-09-938-842A-13/c
; Sequence 13, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 13
; LENGTH: 3501
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

US-09-938-842A-13

Alignment Scores:

Pred. No.:	9.19	Length:	3501
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	9	Gaps:	0

US-09-823-649A-1 (1-11) x US-09-938-842A-13 (1-3501)

```
Qy      1 Leu**************************Glu 11
          |||                           |||
Db    332 CTCTCATCCTCTTCTTCAACGTCATCTTCAGAA 300
```

RESULT 15
US-09-992-598-32
; Sequence 32, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi,Avi J.
; APPLICANT: Baker,Kevin P.
; APPLICANT: Botstein,David
; APPLICANT: Desnoyers,Luc
; APPLICANT: Eaton,Dan L.
; APPLICANT: Ferrara,Napoleone
; APPLICANT: Fong,Sherman
; APPLICANT: Gerber,Hanspeter
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,J.Christopher
; APPLICANT: Gurney,Austin L.
; APPLICANT: Kljavin,Ivar J.
; APPLICANT: Napier,Mary A.
; APPLICANT: Pan,James
; APPLICANT: Paoni,Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	9.22	Length:	3531
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	9	Gaps:	0

US-09-823-649A-1 (1-11) x US-09-992-598-32 (1-3531)

```
Qy       1 Leu**************************Glu 11
             |||                          |||
Db      74 CTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAG 106
```

Search completed: January 21, 2003, 10:07:39
Job time : 31.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds
(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649A-1
Perfect score: 18
Sequence: 1 LXXXXXXXXXE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24616/app_query.fasta_1.1393
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649_@CGN_1_1_2820_@runat_21012003_093151_24616 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	9	50.0	110	13	BM289398	BM289398 Sequence
c	2	9	50.0	112	12	BE842522	BE842522 IL3-ST014
c	3	9	50.0	112	12	BE842524	BE842524 IL3-ST014
c	4	9	50.0	113	10	AW389253	AW389253 IL3-ST014
c	5	9	50.0	113	12	BE842537	BE842537 IL3-ST014
c	6	9	50.0	114	12	BE842532	BE842532 IL3-ST014
c	7	9	50.0	114	12	BE842534	BE842534 IL3-ST014
c	8	9	50.0	127	17	AQ427805	AQ427805 CITBI-E1-
c	9	9	50.0	132	10	AW389255	AW389255 IL3-ST014
	10	9	50.0	147	14	BQ636167	BQ636167 hd05d05.y
	11	9	50.0	151	14	D28324	D28324 RICR3149A R
c	12	9	50.0	157	12	BE842527	BE842527 IL3-ST014
c	13	9	50.0	170	9	AA383829	AA383829 EST97265
	14	9	50.0	181	12	BE842526	BE842526 IL3-ST014
	15	9	50.0	193	12	BE842523	BE842523 IL3-ST014
	16	9	50.0	203	13	BI903789	BI903789 603167684
c	17	9	50.0	230	12	BF803350	BF803350 QV1-CI017
c	18	9	50.0	237	12	BF883878	BF883878 QV1-ET018
	19	9	50.0	240	14	C46922	C46922 C46922 Yuji
c	20	9	50.0	247	12	BG467560	BG467560 Na_L4_04C
	21	9	50.0	249	9	AU076866	AU076866 AU076866
c	22	9	50.0	250	12	BG468040	BG468040 Na_L4_02H
c	23	9	50.0	254	12	BG467571	BG467571 Na_L4_04D
	24	9	50.0	256	14	BQ777127	BQ777127 i143g01.y
	25	9	50.0	259	14	BQ978708	BQ978708 QH16A16.y
	26	9	50.0	267	10	AV869931	AV869931 AV869931
	27	9	50.0	270	17	CNS07FIB	AL608501 Anopheles
	28	9	50.0	273	13	BM138560	BM138560 WHE0494_B
c	29	9	50.0	278	9	AL134269	AL134269 DKFZp547G
c	30	9	50.0	281	14	BQ162056	BQ162056 WHE0494_B
	31	9	50.0	287	10	BB567353	BB567353 BB567353
	32	9	50.0	294	10	BB241459	BB241459 BB241459
c	33	9	50.0	297	10	AV981681	AV981681 AV981681
	34	9	50.0	300	9	AJ461594	AJ461594 AJ461594
	35	9	50.0	300	9	AU099432	AU099432 AU099432
	36	9	50.0	300	14	C11627	C11627 C11627 Yuji
	37	9	50.0	300	14	C31277	C31277 C31277 Yuji
	38	9	50.0	300	14	C55118	C55118 C55118 Yuji
	39	9	50.0	300	14	C55375	C55375 C55375 Yuji
c	40	9	50.0	301	10	AW610337	AW610337 RC4-ST031
c	41	9	50.0	303	9	AI370697	AI370697 ta53a08.x
	42	9	50.0	310	10	AV947389	AV947389 AV947389
	43	9	50.0	321	13	BM280158	BM280158 zah7073 Z
	44	9	50.0	322	14	BM755663	BM755663 K-EST0033
c	45	9	50.0	324	13	BM497825	BM497825 952030C12

ALIGNMENTS

RESULT 1
BM289398
LOCUS BM289398 110 bp mRNA linear EST 29-MAY-2002
DEFINITION Sequence 5 Oryza sativa Nootripathu Differential Display Oryza sativa cDNA, mRNA sequence.
ACCESSION BM289398
VERSION BM289398.1 GI:21245160
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 110)
AUTHORS Raja,J.A.J., Raveendran,M. and Sadasivam,S.

```
  TITLE     Differentially Expressed Sequence Tags in the roots of drought
            tolerant rice variety Nootripathu during water stress
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Raja, J. A. J.; Raveendran, M.; Sadasivam, S.
            Centre for Plant Molecular Biology
            Tamil Nadu Agricultural University
            Coimbatore - 641003, India
            Email: jajraja@yahoo.com.
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Oryza sativa"
                     /cultivar="Nootripathu"
                     /db_xref="taxon:4530"
                     /clone_lib="Oryza sativa Nootripathu Differential Display"
                     /tissue_type="Root tissue"
                     /dev_stage="Vegetative"
                     /note="Differentially expressing mRNAs were identified by
                     differential display of total RNA isolated from the roots
                     of normal and water stressed (70% leaf relative water
                     content) plants. The upregulation of the ESTs during
                     drought was confirmed by northern hybridization analyses."
BASE COUNT       27 a     27 c     20 g     35 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.18e-05        Length:         110
Score:                  9.00            Matches:        2
Percent Similarity:     18.18%          Conservative:   0
Best Local Similarity:  18.18%          Mismatches:     9
Query Match:            50.00%          Indels:         0
DB:                     13              Gaps:           0

US-09-823-649A-1 (1-11) x BM289398 (1-110)

Qy       1 Leu**************************Glu 11
           |||                           |||
Db       8 CTTACCGCAACTACAACGACCGCCACTACGGAG 40

RESULT 2
BE842522/c
LOCUS       BE842522                 112 bp    mRNA    linear   EST 22-SEP-2000
DEFINITION  IL3-ST0142-070600-167-A10 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE842522
VERSION     BE842522.1  GI:10274900
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 112)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=IL3-ST0142-070
            600-167-A10&t3=2000-06-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 112.
FEATURES             Location/Qualifiers
     source          1. .112
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0142"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       13 a     45 c     34 g     20 t
ORIGIN

Alignment Scores:
Pred. No.:              2.22e-05        Length:         112
Score:                  9.00            Matches:        2
Percent Similarity:     18.18%          Conservative:   0
Best Local Similarity:  18.18%          Mismatches:     9
Query Match:            50.00%          Indels:         0
DB:                     12              Gaps:           0

US-09-823-649A-1 (1-11) x BE842522 (1-112)

Qy       1 Leu**************************Glu 11
           |||                           |||
Db      34 TTAGCAGCAGCAGCAGCAGCAGCATCGGAG 2

RESULT 3
BE842524/c
LOCUS       BE842524                 112 bp    mRNA    linear   EST 22-SEP-2000
DEFINITION  IL3-ST0142-070600-167-C04 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE842524
VERSION     BE842524.1  GI:10274902
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 112)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=IL3-ST0142-070
            600-167-C04&t3=2000-06-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 112.
FEATURES             Location/Qualifiers
     source          1. .112
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0142"
                     /dev_stage="Adult"
```

/note="Organ: stomach; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 13 a 46 c 34 g 19 t
ORIGIN

Alignment Scores:

Pred. No.:	2.22e-05	Length:	112
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	12	Gaps:	0

US-09-823-649A-1 (1-11) x BE842524 (1-112)

```
Qy       1 Leu**************************Glu 11
             |||                          |||
Db      34 TTAGCAGCAGCAGCAGCAGCGGCATCGGAG 2
```

RESULT 4
AW389253/c
LOCUS AW389253 113 bp mRNA linear EST 04-FEB-2000
DEFINITION IL3-ST0142-131099-018-C02 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW389253
VERSION AW389253.1 GI:6893912
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 113)
AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-ST0142-131099-018-C02&t3=1999-10-13&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 27
High quality sequence stop: 113.
FEATURES Location/Qualifiers
source 1. .113
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ST0142"
/dev_stage="Adult"
/note="Organ: stomach; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 14 a 41 c 33 g 25 t
ORIGIN

Alignment Scores:

Pred. No.:	2.24e-05	Length:	113
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-1 (1-11) x AW389253 (1-113)

```
Qy       1 Leu**************************Glu 11
             |||                          |||
Db      35 TTAGCAGCAGCAACAGCAGCAGCAGCATCGGAA 3
```

RESULT 5
BE842537/c
LOCUS BE842537 113 bp mRNA linear EST 22-SEP-2000
DEFINITION IL3-ST0142-120600-174-E05 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE842537
VERSION BE842537.1 GI:10274915
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 113)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL3-ST0142-120600-174-E05&t3=2000-06-12&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 113.
FEATURES Location/Qualifiers
source 1. .113
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ST0142"
/dev_stage="Adult"
/note="Organ: stomach; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 14 a 45 c 34 g 20 t
ORIGIN

Alignment Scores:

Pred. No.:	2.24e-05	Length:	113
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	12	Gaps:	0

US-09-823-649A-1 (1-11) x BE842537 (1-113)

```
Qy      1 Leu**************************Glu 11
          |||                           |||
Db     35 TTAGCAGCAGCAGCAGCAGCAGCATCGGAG 3

RESULT 6
BE842532/c
LOCUS       BE842532                114 bp    mRNA    linear   EST 22-SEP-2000
DEFINITION  IL3-ST0142-070600-168-F05 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE842532
VERSION     BE842532.1  GI:10274910
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 114)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1-&t2-IL3-ST0142-070
            600-168-F05&t3=2000-06-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 114.
FEATURES             Location/Qualifiers
     source          1. .114
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0142"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: pucl8; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       14 a     46 c     34 g     20 t
ORIGIN

Alignment Scores:
Pred. No.:              2.26e-05       Length:         114
Score:                  9.00           Matches:        2
Percent Similarity:     18.18%         Conservative:   0
Best Local Similarity:  18.18%         Mismatches:     9
Query Match:            50.00%         Indels:         0
DB:                     12             Gaps:           0

US-09-823-649A-1 (1-11) x BE842532 (1-114)

Qy      1 Leu**************************Glu 11
          |||                           |||
Db     36 TTAGCAGCAGCAGCAGCAGCGGCATCGGAG 4

RESULT 7
BE842534/c
LOCUS       BE842534                114 bp    mRNA    linear   EST 22-SEP-2000
DEFINITION  IL3-ST0142-070600-168-G05 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE842534
VERSION     BE842534.1  GI:10274912
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 114)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1-&t2-IL3-ST0142-070
            600-168-G05&t3-2000-06-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 114.
FEATURES             Location/Qualifiers
     source          1. .114
                     /organism-"Homo sapiens"
                     /db_xref-"taxon:9606"
                     /clone_lib-"ST0142"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: pucl8; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       14 a     46 c     34 g     20 t
ORIGIN

Alignment Scores:
Pred. No.:              2.26e-05       Length:         114
Score:                  9.00           Matches:        2
Percent Similarity:     18.18%         Conservative:   0
Best Local Similarity:  18.18%         Mismatches:     9
Query Match:            50.00%         Indels:         0
DB:                     12             Gaps:           0

US-09-823-649A-1 (1-11) x BE842534 (1-114)

Qy      1 Leu**************************Glu 11
          |||                           |||
Db     36 TTAGCAGCAGCAGCAGCAGCGGCATCGGAG 4

RESULT 8
AQ427805/c
LOCUS       AQ427805                127 bp    DNA     linear   GSS 24-MAR-1999
DEFINITION  CITBI-E1-2562E1.TF CITBI-E1 Homo sapiens genomic clone 2562E1, DNA
            sequence.
ACCESSION   AQ427805
VERSION     AQ427805.1  GI:4500712
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 127)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CITBI-E1-2562E1.TR
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .127
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2562E1"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"
BASE COUNT       34 a     22 c     25 g     46 t
ORIGIN

Alignment Scores:
Pred. No.:              2.51e-05        Length:         127
Score:                  9.00            Matches:        2
Percent Similarity:     18.18%          Conservative:   0
Best Local Similarity:  18.18%          Mismatches:     9
Query Match:            50.00%          Indels:         0
DB:                     17              Gaps:           0

US-09-823-649A-1 (1-11) x AQ427805 (1-127)

Qy       1 Leu**************************Glu 11
             |||                          |||
Db      59 CTTACATCAAGTGCAGCATCCTCATCATCTGAA 27

RESULT 9
AW389255/c
LOCUS       AW389255                 132 bp    mRNA    linear   EST 04-FEB-2000
DEFINITION  IL3-ST0142-131099-018-E01 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW389255
VERSION     AW389255.1  GI:6893914
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 132)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=IL3&t2=IL3-ST0142-
            131099-018-E01&t3=1999-10-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 132.
FEATURES             Location/Qualifiers
     source          1. .132
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0142"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: pucl8; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       21 a     33 c     58 g     20 t
ORIGIN

Alignment Scores:
Pred. No.:              2.61e-05        Length:         132
Score:                  9.00            Matches:        2
Percent Similarity:     18.18%          Conservative:   0
Best Local Similarity:  18.18%          Mismatches:     9
Query Match:            50.00%          Indels:         0
DB:                     10              Gaps:           0

US-09-823-649A-1 (1-11) x AW389255 (1-132)

Qy       1 Leu**************************Glu 11
             |||                          |||
Db     115 TTAGCAGCAGCAGCAGCAGCAGCATCGGAG 83

RESULT 10
BQ636167
LOCUS       BQ636167                 147 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  hd05d05.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
            Homo sapiens cDNA clone hd05d05 5', mRNA sequence.
ACCESSION   BQ636167
VERSION     BQ636167.1  GI:21760626
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 147)
  AUTHORS   Wistow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman
            ,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of human retina for the NEIBank
            Project: Retbindin, an abundant, novel retinal cDNA and alternative
            splicing of other retina-preferred gene transcripts
  JOURNAL   Mol. Vis. 8 (4), (2002) In press
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 05  row: d  column: 05
            Seq primer: M13RPl reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .147
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="hd05d05"
                     /clone_lib="Human Retina cDNA (Un-normalized, unamplified
                     ): hd/he"
                     /tissue_type="Retina"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
```

was dissected from two 80 year old donors with no observed eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

BASE COUNT 25 a 50 c 55 g 17 t
ORIGIN

Alignment Scores:

Pred. No.:	2.91e-05	Length:	147
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	14	Gaps:	0

US-09-823-649A-1 (1-11) x BQ636167 (1-147)

```
Qy       1 Leu**************************Glu 11
             |||                          |||
Db      82 TTAGCAGCAGCAGCAGCAGCAGCATCGGAG 114
```

RESULT 11
D28324
LOCUS D28324 151 bp mRNA linear EST 03-APR-2002
DEFINITION RICR3149A Rice root Oryza sativa (japonica cultivar-group) cDNA clone R3149_1A, mRNA sequence.
ACCESSION D28324
VERSION D28324.1 GI:454555
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group).
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 151)
AUTHORS Minobe,Y. and Sasaki,T.
TITLE Rice cDNA from root
JOURNAL Unpublished (1995)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT ='RGP'.

FEATURES	Location/Qualifiers
source	1. .151
	/organism="Oryza sativa (japonica cultivar-group)"
	/cultivar="Nipponbare"
	/db_xref="taxon:39947"
	/clone="R3149_1A"
	/clone_lib="Rice root"
	/note="Prepared from seedling root. "

BASE COUNT 19 a 52 c 55 g 20 t 5 others
ORIGIN

Alignment Scores:

Pred. No.:	2.98e-05	Length:	151
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	14	Gaps:	0

US-09-823-649A-1 (1-11) x D28324 (1-151)

```
Qy       1 Leu**************************Glu 11
             |||                          |||
Db      19 CTCGCCGCCGCCGCCGCCGCCGCATCTGCCGAA 51
```

RESULT 12
BE842527/c
LOCUS BE842527 157 bp mRNA linear EST 22-SEP-2000
DEFINITION IL3-ST0142-070600-167-F08 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE842527
VERSION BE842527.1 GI:10274905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 157)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/qethtml2.pl?t1-&t2-IL3-ST0142-070600-167-F08&t3-2000-06-07&t4-1)
Seq primer: puc 18 forward
High quality sequence stop: 157.

FEATURES	Location/Qualifiers
source	1. .157
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone_lib="ST0142"
	/dev_stage="Adult"
	/note="Organ: stomach; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A size library was made by cloning products derived from cDNA/ESTs PCR (U.S. Letters Patent application No. 196,716 Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 20 a 56 c 57 g 24 t
ORIGIN

Alignment Scores:

Pred. No.:	3.1e-05	Length:	157
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	12	Gaps:	0

US-09-823-649A-1 (1-11) x BE842527 (1-157)

```
Qy       1 Leu**************************Glu 11
             |||                          |||
Db      79 TTAGCAGCAGCAGCAGCAGCAGCATCGGAG 47
```

RESULT 13

```
AA383829/c
LOCUS       AA383829                170 bp    mRNA    linear   EST 21-APR-1997
DEFINITION  EST97265 Testis I Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION   AA383829
VERSION     AA383829.1  GI:2036147
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 170)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
            Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .170
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):188010"
                     /db_xref="taxon:9606"
                     /clone_lib="Testis I"
                     /sex="male"
                     /dev_stage="adult"
                     /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
BASE COUNT       36 a     71 c     32 g     29 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              3.36e-05       Length:         170
Score:                  9.00           Matches:        2
Percent Similarity:     18.18%         Conservative:   0
Best Local Similarity:  18.18%         Mismatches:     9
Query Match:            50.00%         Indels:         0
DB:                     9              Gaps:           0

US-09-823-649A-1 (1-11) x AA383829 (1-170)

Qy      1 Leu**************************Glu 11
          |||                           |||
Db    135 CTTAGCGCTGCTGCGGCTGCCGCAGCTGCCGAA 103

RESULT 14
BE842526
LOCUS       BE842526                181 bp    mRNA    linear   EST 22-SEP-2000
DEFINITION  IL3-ST0142-070600-167-F04 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE842526
VERSION     BE842526.1  GI:10274904
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 181)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL3-ST0142-070
            600-167-F04&t3=2000-06-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 181.
FEATURES             Location/Qualifiers
     source          1. .181
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0142"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       29 a     58 c     65 g     29 t
ORIGIN

Alignment Scores:
Pred. No.:              3.57e-05       Length:         181
Score:                  9.00           Matches:        2
Percent Similarity:     18.18%         Conservative:   0
Best Local Similarity:  18.18%         Mismatches:     9
Query Match:            50.00%         Indels:         0
DB:                     12             Gaps:           0

US-09-823-649A-1 (1-11) x BE842526 (1-181)

Qy      1 Leu**************************Glu 11
          |||                           |||
Db     79 TTAGCAGCAGCAGCAGCAGCGGCATCGGAG 111

RESULT 15
BE842523
LOCUS       BE842523                193 bp    mRNA    linear   EST 22-SEP-2000
DEFINITION  IL3-ST0142-070600-167-B06 ST0142 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE842523
VERSION     BE842523.1  GI:10274901
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 193)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel. +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL3-ST0142-070
            600-167-B06&t3=2000-06-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 193.
FEATURES             Location/Qualifiers
     source          1. .193
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ST0142"
                     /dev_stage="Adult"
                     /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       26 a     76 c     61 g     30 t
ORIGIN

Alignment Scores:
Pred. No.:              3.81e-05        Length:         193
Score:                  9.00            Matches:        2
Percent Similarity:     18.18%          Conservative:   0
Best Local Similarity:  18.18%          Mismatches:     9
Query Match:            50.00%          Indels:         0
DB:                     12              Gaps:           0

US-09-823-649A-1 (1-11) x BE842523 (1-193)

Qy       1 Leu**************************Glu 11
             |||                          |||
Db      79 TTAGCAGCAGCAGCAGCAGCAGCATCGGAG 111


Search completed: January 21, 2003, 12:24:10
Job time : 1171.14 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 1168.14 Seconds
(without alignments)
152.507 Million cell updates/sec

Title: US-09-823-649A-7
Perfect score: 49
Sequence: 1 LAQNLNIXRKE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: ..
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24616/app_query.fasta_1.1393
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649_@CGN_1_1_2820_@runat_21012003_093151_24616 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	42	85.7	695	10	BB468375	BB468375 BB468375
2	40	81.6	359	13	BI495153	BI495153 df115d10.
3	40	81.6	501	13	BI492651	BI492651 d127c10.w
4	39	79.6	832	17	AQ855611	AQ855611 CpG1705A
5	38	77.6	449	12	BF635786	BF635786 NF039G09D
6	38	77.6	1034	17	AC155700	AC155700 Pan trogl
7	37	75.5	175	12	BF371247	BF371247 RC6-BN013
c 8	37	75.5	187	13	BM026044	BM026044 fu77e06.x
c 9	37	75.5	261	9	AV245515	AV245515 AV245515
10	37	75.5	427	14	BQ612936	BQ612936 sap78h07.
11	37	75.5	520	17	AQ732941	AQ732941 HS_5548_B
12	37	75.5	528	17	AQ708467	AQ708467 HS_5221_B
13	37	75.5	675	17	AC178233	AC178233 Pan trogl
c 14	37	75.5	718	17	AC172746	AC172746 Pan trogl
15	37	75.5	738	10	AV709998	AV709998 AV709998
c 16	37	75.5	753	12	BE821761	BE821761 GM700015A
c 17	37	75.5	756	10	AW350371	AW350371 GM210008A
c 18	37	75.5	761	17	AQ745123	AQ745123 HS_5501_A
c 19	37	75.5	819	17	BH248796	BH248796 BOGAU79TF
c 20	37	75.5	838	14	BQ441565	BQ441565 AGENCOURT
21	37	75.5	888	17	BH158636	*BH158636 ENTTM53TR*
c 22	37	75.5	1084	17	BH133865	BH133865 ENTOB33TF
c 23	37	75.5	1155	17	AG039702	AG039702 Pan trogl
24	36	73.5	271	13	BI681170	BI681170 460478 MA
c 25	36	73.5	378	17	AZ391764	AZ391764 1M0154O06
c 26	36	73.5	401	17	BH754534	BH754534 SALK_0425
c 27	36	73.5	431	17	BH754529	BH754529 SALK_0425
28	36	73.5	498	9	AU084303	AU084303 AU084303
29	36	73.5	554	17	AQ727741	AQ727741 HS_5457_A
c 30	36	73.5	618	13	BJ070175	BJ070175 BJ070175
31	36	73.5	620	14	BQ603915	BQ603915 MI-P-CP1-
32	36	73.5	629	10	AW616886	AW616886 EST323297
c 33	36	73.5	665	17	AQ316963	AQ316963 CITBI-E1-
34	36	73.5	680	10	BB639624	BB639624 BB639624
c 35	36	73.5	698	12	BE891760	BE891760 601434485
c 36	36	73.5	740	17	AQ316982	AQ316982 CITBI-E1-
c 37	36	73.5	745	17	AQ308677	AQ308677 CITBI-E1-
c 38	36	73.5	748	17	AQ330811	AQ330811 nbxb0047P
39	36	73.5	938	14	BQ689597	BQ689597 AGENCOURT
c 40	36	73.5	944	17	AZ207883	AZ207883 SP_0135_A
41	36	73.5	974	12	BG026212	BG026212 602291545
42	36	73.5	1037	17	CNS02Q8V	AL208984 Tetraodon
43	36	73.5	1094	14	BM811021	BM811021 AGENCOURT
44	35	71.4	111	13	BJ386485	BJ386485 BJ386485
45	35	71.4	164	9	AI318539	AI318539 ta74106.x

ALIGNMENTS

RESULT 1
BB468375/c
LOCUS BB468375 695 bp mRNA linear EST 25-OCT-2001
DEFINITION BB468375 RIKEN full-length enriched, 12 days embryo eyeball Mus musculus cDNA clone D230019D05 3', mRNA sequence.
ACCESSION BB468375
VERSION BB468375.2 GI:16427508
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 695)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda

,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

JOURNAL Unpublished (2001)

COMMENT On Jul 22, 2000 this sequence version replaced gi:9385564.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)

wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara ,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)

Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES Location/Qualifiers
source 1. .695
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="D230019D05"
/clone_lib="RIKEN full-length enriched, 12 days embryo eyeball"
/tissue_type="eyeball"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 163 a 148 c 114 g 267 t 3 others

ORIGIN

Alignment Scores:

Pred. No.:	42.1	Length:	695
Score:	42.00	Matches:	9
Percent Similarity:	90.00%	Conservative:	0
Best Local Similarity:	90.00%	Mismatches:	1
Query Match:	85.71%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-7 (1-11) x BB468375 (1-695)

```
Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLys 10
          ||||||||||||||||||||   ||||||
Db    337 TTAGCTCAAAACCTAAACATAACGAGAAAA 308
```

RESULT 2
BI495153
LOCUS BI495153 359 bp mRNA linear EST 28-AUG-2001
DEFINITION dfl15d10.w1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:2539458 3', mRNA sequence.
ACCESSION BI495153
VERSION BI495153.1 GI:15334497
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R. and Morton,C.C.
TITLE Isolation of novel and known genes from a human fetal cochlear cDNA library using subtractive hybridization and differential screening
JOURNAL Genomics 23, 42-50 (1994)
MEDLINE 95130111
COMMENT Contact: Morton, C. C.
Departments of Pathology and Obstetrics, Gynecology and Reproductive Biology
Brigham and Women's Hospital
75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
Tel: 617 732 7980
Fax: 617 738 6996
Email: ccmorton@bics.bwh.harvard.edu
DNA sequencing and analyses were performed by National Institutes of Health Intramural Sequencing Center (NISC; see http://www.nisc.nih.gov).
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Plate: LLAM6325 row: H column: 19
Seq primer: T7 primer.

FEATURES Location/Qualifiers
source 1. .359
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2539458"
/clone_lib="Morton Fetal Cochlea"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned unidirectionally. Primer: Oligo dT. Fetal cochlea, normal. 37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP XR Vector. Library constructed by N. Robertson, C. Morton. ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

BASE COUNT 102 a 80 c 50 g 127 t

ORIGIN

Alignment Scores:

Pred. No.:	45.4	Length:	359
Score:	40.00	Matches:	8
Percent Similarity:	90.00%	Conservative:	1
Best Local Similarity:	80.00%	Mismatches:	1

```
Query Match:            81.63%      Indels:        0
DB:                     13          Gaps:          0

US-09-823-649A-7 (1-11) x BI495153 (1-359)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLys 10
          ||||||||||||||||||:::   ||||||
Db    266 TTGGCCCAAAACCTCAACTTAAAAAGGAAA 295

RESULT 3
BI492651
LOCUS       BI492651                 501 bp    mRNA    linear   EST 28-AUG-2001
DEFINITION  df27c10.w1 Morton Fetal Cochlea Homo sapiens cDNA clone
            IMAGE:2484498 3', mRNA sequence.
ACCESSION   BI492651
VERSION     BI492651.1  GI:15331995
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.
            and Morton,C.C.
  TITLE     Isolation of novel and known genes from a human fetal cochlear cDNA
            library using subtractive hybridization and differential screening
  JOURNAL   Genomics 23, 42-50 (1994)
  MEDLINE   95130111
COMMENT     Contact: Morton, C. C.
            Departments of Pathology and Obstetrics, Gynecology and
            Reproductive Biology
            Brigham and Women's Hospital
            75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
            Tel: 617 732 7980
            Fax: 617 738 6996
            Email: ccmorton@bics.bwh.harvard.edu
            DNA sequencing and analyses were performed by National Institutes
            of Health Intramural Sequencing Center (NISC; see
            http://www.nisc.nih.gov).
             This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Plate: LLAM6182  row: F  column: 19
            Seq primer: T7 primer.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2484498"
                     /clone_lib="Morton Fetal Cochlea"
                     /tissue_type="cochlea"
                     /dev_stage="16-22 week fetus"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI;
                     Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
                     unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
                     37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
                     XR Vector. Library constructed by N. Robertson, C. Morton.
                     -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      145 a    110 c     81 g    165 t
ORIGIN

Alignment Scores:
Pred. No.:              73          Length:        501
Score:                  40.00       Matches:       8
Percent Similarity:     90.00%      Conservative:  1
Best Local Similarity:  80.00%      Mismatches:    1
Query Match:            81.63%      Indels:        0
DB:                     13          Gaps:          0

US-09-823-649A-7 (1-11) x BI492651 (1-501)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLys 10
          ||||||||||||||||||:::   ||||||
Db    263 TTGGCCCAAAACCTTAACTTAAAAAGGAAA 292

RESULT 4
AQ855611
LOCUS       AQ855611                 832 bp    DNA     linear   GSS 12-MAY-2000
DEFINITION  CpG1705A CpIOWAgDNA1 Cryptosporidium parvum genomic, DNA sequence.
ACCESSION   AQ855611
VERSION     AQ855611.1  GI:6206068
KEYWORDS    GSS.
SOURCE      Cryptosporidium parvum.
  ORGANISM  Cryptosporidium parvum
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Cryptosporidiidae; Cryptosporidium.
REFERENCE   1  (bases 1 to 832)
  AUTHORS   Strong,W.B. and Nelson,R.G.
  TITLE     Preliminary profile of the Cryptosporidium parvum genome: an
            expressed sequence tag and genome survey sequence analysis
  JOURNAL   Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)
  MEDLINE   20183851
COMMENT     Contact: Nelson, R. G.
            Depts. of Medicine & Pharmaceutical Chemistry
            San Francisco General Hospital-University of California, San
            Francisco
            Box 0811, San Francisco, CA 94143-0811, USA
            Tel: 415 206 8846
            Fax: 415 206 3353
            Email: malaria@itsa.ucsf.edu
            For Annotation Data see http://medsfgh.ucsf.edu/id/CpTags/home.html
            Seq primer: M13(-21) forward
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .832
                     /organism="Cryptosporidium parvum"
                     /strain="IOWA"
                     /db_xref="taxon:5807"
                     /clone_lib="CpIOWAgDNA1"
                     /lab_host="E. coli XL2 Blue MRF'"
                     /note="Vector: pBlueScript II (SK-); Site_1: EcoRV; C.
                     parvum (IOWA isolate) genomic DNA was hydrodynamically
                     sheared  to produce fragments having a tight size
                     distribution between 2-4 kb by Dr. Yvonne Thorstenson of
                     the Stanford DNA Sequencing and Technology Center
                     (http://sequence-www.stanford.edu/group/techdev/shear.htm
                     ).  The randomly sheared gDNA was chromatographed on
                     Sephacryl S-400 to remove any small fragments and DNA
                     eluting in the void volume was subcloned into an EcoR
                     V-digested, alkaline phosphatase-treated pBlueScript II
                     (SK-) vector and transformed into E. coli strain XL2 Blue
                     MRF'.  Recombinant clones from the first plating of the
                     library were selected for sequence analysis using T3 and
                     T7 primers."
BASE COUNT      335 a    114 c    196 g    185 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              250         Length:        832
Score:                  39.00       Matches:       7
Percent Similarity:     90.00%      Conservative:  2
Best Local Similarity:  70.00%      Mismatches:    1
Query Match:            79.59%      Indels:        0
DB:                     17          Gaps:          0

US-09-823-649A-7 (1-11) x AQ855611 (1-832)

Qy      2 AlaGlnAsnLeuAsnIle***ArgLysGlu 11
          ||||||||||||||||:::   :::||||||
Db    626 GCTCAGAACTTGAACTTAACGAAAAAAGAA 655

RESULT 5
BF635786
LOCUS       BF635786                 449 bp    mRNA    linear   EST 19-DEC-2000
DEFINITION  NF039G09DT1F1069 Drought Medicago truncatula cDNA clone NF039G09DT
```

```
          5', mRNA sequence.
ACCESSION   BF635786
VERSION     BF635786.1  GI:11899944
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Torrez-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula drought library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7391
            Fax: 580 221 7380
            Email: gdmay@noble.org
            Insert Length: 449   Std Error: 0.00
            Plate: 039  row: G  column: 09
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .449
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF039G09DT"
                     /clone_lib="Drought"
                     /tissue_type="Plantlets"
                     /dev_stage="Pooled timepoints"
                     /note="Vector: Lambda Zap; Contains a mixture of entire
                     plantlets harvested in a series of days-post-watering
                     timepoints. "
BASE COUNT      126 a     74 c    118 g    131 t
ORIGIN

Alignment Scores:
Pred. No.:              172           Length:         449
Score:                  38.00         Matches:        8
Percent Similarity:     88.89%        Conservative:   0
Best Local Similarity:  88.89%        Mismatches:     1
Query Match:            77.55%        Indels:         0
DB:                     12            Gaps:           0

US-09-823-649A-7 (1-11) x BF635786 (1-449)

Qy      1 LeuAlaGlnAsnLeuAsnIle***Arg 9
          |||||||||||||||||||||   |||
Db    347 TTGGCCCAAAATTTGAACATCACGAGG 373

RESULT 6
AG155700
LOCUS       AG155700               1034 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-020B14.TJ, genomic survey
            sequence.
ACCESSION   AG155700
VERSION     AG155700.1  GI:16685378
KEYWORDS    GSS.
SOURCE      Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
            Male BAC Library clone:RP43-020B14.TJ.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1034)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: TJ
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1034
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="RP43-020B14.TJ"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      306 a    202 c    282 g    235 t      9 others
ORIGIN

Alignment Scores:
Pred. No.:              567           Length:         1034
Score:                  38.00         Matches:        8
Percent Similarity:     88.89%        Conservative:   0
Best Local Similarity:  88.89%        Mismatches:     1
Query Match:            77.55%        Indels:         0
DB:                     17            Gaps:           0

US-09-823-649A-7 (1-11) x AG155700 (1-1034)

Qy      2 AlaGlnAsnLeuAsnIle***ArgLys 10
            ||||||||||||||||||   ||||||
Db    468 GCACAAAATTTAAACATTAAACGAAAA 494

RESULT 7
BF371247
LOCUS       BF371247               175 bp    mRNA    linear   EST 24-NOV-2000
DEFINITION  RC6-FN0138-260700-011-A01 FN0138 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF371247
VERSION     BF371247.1  GI:11333876
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 175)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
```

Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC6&t2=RC6-FN0138-260700-011-A01&t3=2000-07-26&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 175.

FEATURES Location/Qualifiers
source 1. .175
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FN0138"
/dev_stage="Adult"
/note="Organ: prostate_normal; Vector: puc18; Site_1: SmaI ; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 38 a 48 c 28 g 61 t
ORIGIN

Alignment Scores:

Pred. No.:	74.7	Length:	175
Score:	37.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	75.51%	Indels:	0
DB:	12	Gaps:	0

US-09-823-649A-7 (1-11) x BF371247 (1-175)

```
Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          |||   ||||||||||||||||   :::|||:::
Db     50 CTAGTTCAGAACCTTAATATCTCCCAGAAGGAT 82
```

RESULT 8
BM026044/c
LOCUS BM026044 187 bp mRNA linear EST 30-OCT-2001
DEFINITION fu77e06.x1 Gong zebrafish testis Danio rerio cDNA clone 5376875 3' similar to SW:IF36_HUMAN Q64252 EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT 6 ;, mRNA sequence.
ACCESSION BM026044
VERSION BM026044.1 GI:16540400
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 187)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
The library was constructed by Dr. Z. Gong. DNA Sequencing by: Washington University Genome Sequencing Center St. Louis. Please contact Zhiyuan Gong for further information on this library (National University of Singapore, Department of Biological Sciences, Lower Kent Ridge Road, Singapore 119260).
Trace considered overall poor quality
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .187
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5376875"
/clone_lib="Gong zebrafish testis"
/sex-"male"
/dev_stage-"4-5 month"
/lab_host-"DH10B (phage-resistant)"
/note="Organ: testis (pooled); Vector: pBluescript SK-; Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolatd from the testes of 31 male adult zebrafish (4-5 month old). cDNAs were made using oligo-dT primers and inserted into lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo mass-excised to pBluescript SK- following the Washington University protocol (http://genome.wustl.edu/est/lambda_protocol.shtml). Please contact Zhiyuan Gong for further information on this library (National University of Singapore, Department of Biological Sciences, Lower Kent Ridge Road, Singapore 119260)."

BASE COUNT 51 a 46 c 42 g 48 t
ORIGIN

Alignment Scores:

Pred. No.:	82.1	Length:	187
Score:	37.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	75.51%	Indels:	0
DB:	13	Gaps:	0

US-09-823-649A-7 (1-11) x BM026044 (1-187)

```
Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          :::|||   ||||||||||||   |||:::|||
Db     70 ATGGCGGATAACCTGAATATATCTCGTGAGGAG 38
```

RESULT 9
AV245515/c
LOCUS AV245515 261 bp mRNA linear EST 04-NOV-1999
DEFINITION AV245515 RIKEN full-length enriched, 0 day neonate head Mus musculus cDNA clone 4831440J19 3' similar to D50263 Mus musculus DAN mRNA, mRNA sequence.
ACCESSION AV245515
VERSION AV245515.1 GI:6232974
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 261)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai ,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata ,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Takahashi,F., Tateno,M., Tominaga,N., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al. 1999)
JOURNAL Unpublished (1999)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,

```
            Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki
            ,Y.
             Transcriptional sequencing: A method for DNA sequencing using RNA
            polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .261
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4831440J19"
                     /clone_lib="RIKEN full-length enriched, 0 day neonate
                     head"
                     /sex="mixed"
                     /tissue_type="head"
                     /dev_stage="0 day neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 100.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cloned into the XhoI and BamHI sites.
                     Vector: a modified pBluescript KS(+) after bulk excision
                     from Lambda FLC I."
BASE COUNT       51 a     72 c     51 g     87 t
ORIGIN

Alignment Scores:
Pred. No.:              132          Length:         261
Score:                  37.00        Matches:        8
Percent Similarity:     72.73%       Conservative:   0
Best Local Similarity:  72.73%       Mismatches:     3
Query Match:            75.51%       Indels:         0
DB:                     9            Gaps:           0

US-09-823-649A-7 (1-11) x AV245515 (1-261)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          ||||||||||   ||||||      |||||||||
Db     70 CTAGCCCAAGGGTTAAATTATACTAGGAAAGAA 38

RESULT 10
BQ612936
LOCUS       BQ612936                427 bp    mRNA    linear   EST 26-JUN-2002
DEFINITION  sap78h07.y1 Gm-c1087 Glycine max cDNA clone SOYBEAN CLONE ID:
            Gm-c1087-6901 5' similar to SW:DAPF_SYNY3 P74667 DIAMINOPIMELATE
            EPIMERASE ;, mRNA sequence.
ACCESSION   BQ612936
VERSION     BQ612936.1  GI:21602605
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 427)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact: ccu@resgen.com web site:
            www.resgen.com
            Putative full length read
            vector to vector length is 471
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .427
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="SOYBEAN CLONE ID: Gm-c1087-6901"
                     /clone_lib="Gm-c1087"
                     /tissue_type="Roots"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The mRNA was prepared using polyatract mRNA system
                     from PROMEGA. The cDNA was prepared using the STRATAGENE
                     kit. Complementary DNA was synthesized from mRNA using a
                     primer consisting of a poly(dT) sequence with a XhoI
                     restriction site
                     (5'GACAGAGAGAGACAGAGAGAACTAGTCTCGAGTTTTTTTTTTTTTTTT).
                     EcoRI adapters (5'OH-AATTCGGCACGAG and 3'CCGTGCTCp) were
                     ligated to the blunt-ended cDNA fragments followed by XhoI
                     digestion. The cDNA fragments were directionally cloned
                     into the EcoRI XhoI restriction sites of the pBluescript
                     vector. The vector was previously dephosphorylated. The
                     ligated cDNA fragments were transformed into DH10B by
                     electroporation. Library is in LB medium with ampicillin
                     and glycerol 8%. Average insert size: 800 bp."
BASE COUNT      109 a     78 c     113 g     127 t
ORIGIN

Alignment Scores:
Pred. No.:              267          Length:         427
Score:                  37.00        Matches:        8
Percent Similarity:     88.89%       Conservative:   0
Best Local Similarity:  88.89%       Mismatches:     1
Query Match:            75.51%       Indels:         0
DB:                     14           Gaps:           0

US-09-823-649A-7 (1-11) x BQ612936 (1-427)

Qy      1 LeuAlaGlnAsnLeuAsnIle***Arg 9
          |||||||||||||||||||||   |||
Db     63 TTGGCCCAAAATTTGAACATCATGAGG 89

RESULT 11
AQ732941
LOCUS       AQ732941                520 bp    DNA     linear   GSS 15-JUL-1999
DEFINITION  HS_5548_B1_E01_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1124 Col=1 Row=J, DNA sequence.
ACCESSION   AQ732941
VERSION     AQ732941.1  GI:5504493
KEYWORDS    GSS.
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 520)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1124  row: J  column: 1
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 520.
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=1124 Col=1 Row=J"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      107 a    142 c    112 g    155 t      4 others
ORIGIN

Alignment Scores:
Pred. No.:                353          Length:        520
Score:                    37.00        Matches:       7
Percent Similarity:       81.82%       Conservative:  2
Best Local Similarity:    63.64%       Mismatches:    2
Query Match:              75.51%       Indels:        0
DB:                       17           Gaps:          0

US-09-823-649A-7 (1-11) x AQ732941 (1-520)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          |||   |||||||||||||||   :::|||::: 
Db    146 CTAGTTCAGAACCTTAATATCTCCCAGAAGGAT 178

RESULT 12
AQ708467
LOCUS       AQ708467                528 bp    DNA     linear   GSS 07-JUL-1999
DEFINITION  HS_5221_B2_B03_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=797 Col=6 Row=D, DNA sequence.
ACCESSION   AQ708467
VERSION     AQ708467.1  GI:5417893
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 797  row: D  column: 6
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 528.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=797 Col=6 Row=D"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      167 a    109 c     98 g    144 t     10 others
ORIGIN

Alignment Scores:
Pred. No.:                361          Length:        528
Score:                    37.00        Matches:       7
Percent Similarity:       81.82%       Conservative:  2
Best Local Similarity:    63.64%       Mismatches:    2
Query Match:              75.51%       Indels:        0
DB:                       17           Gaps:          0

US-09-823-649A-7 (1-11) x AQ708467 (1-528)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          |||:::|||||||:::|||||      ||||||
Db    333 TTGTCTCAAAATATAAACATAAATAATAAAGAA 365

RESULT 13
AG178233
LOCUS       AG178233                675 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-050C07.TJ, genomic survey
            sequence.
ACCESSION   AG178233
VERSION     AG178233.1  GI:16707913
KEYWORDS    GSS.
SOURCE      Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
            Male BAC Library clone:RP43-050C07.TJ.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 675)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: TJ
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="RP43-050C07.TJ"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      185 a    164 c    166 g    142 t     18 others
ORIGIN

Alignment Scores:
Pred. No.:              512           Length:        675
Score:                  37.00         Matches:       7
Percent Similarity:     90.91%        Conservative:  3
Best Local Similarity:  63.64%        Mismatches:    1
Query Match:            75.51%        Indels:        0
DB:                     17            Gaps:          0

US-09-823-649A-7 (1-11) x AG178233 (1-675)

Qy       1 LeuAl GlnAsnLeuAsnIle***ArgLysGlu 11
           ||||| ||:::||||||:::   |||:::|||
Db     187 TTGGCC AGAGCCTCAATCTCACCAGGCAGGAA 219

RESULT 14
AG172746/c
LOCUS       AG172746                718 bp    DNA     linear   GSS 09-JAN-2002
DEFINITION  Pan troglodytes DNA, clone: RP43-042H23.TJ, genomic survey
            sequence.
ACCESSION   AG172746
VERSION     AG172746.1  GI:16702426
KEYWORDS    GSS.
SOURCE      Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
            Male BAC Library clone:RP43-042H23.TJ.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library RPCI-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 718)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: TJ
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
                 R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .718
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="RP43-042H23.TJ"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      168 a    174 c    125 g    249 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              560           Length:        718
Score:                  37.00         Matches:       7
Percent Similarity:     90.91%        Conservative:  3
Best Local Similarity:  63.64%        Mismatches:    1
Query Match:            75.51%        Indels:        0
DB:                     17            Gaps:          0

US-09-823-649A-7 (1-11) x AG172746 (1-718)

Qy       1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
           :::|||||||:::::|||||   |||||||||
Db     100 GTTGCTCAAAGTATGAATATTCAAAGGAAAGAA 68

RESULT 15
AV709998
LOCUS       AV709998                738 bp    mRNA    linear   EST 09-OCT-2000
DEFINITION  AV709998 Cu Homo sapiens cDNA clone CuAALH05 5', mRNA sequence.
ACCESSION   AV709998
VERSION     AV709998.1  GI:10728284
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 738)
  AUTHORS   Peng,Y., Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
            Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu
            ,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J.,
            Chen,Z. and Han,Z.
  TITLE     Homo sapiens cDNA Cu clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .738
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CuAALH05"
                     /clone_lib="Cu"
                     /tissue_type="adrenal cortico adenoma for Cushing's
                     syndrome"
                     /dev_stage="Adult"
                     /note="Vector: pBluescript sk(-)"
BASE COUNT      245 a    134 c    143 g    216 t
ORIGIN

Alignment Scores:
Pred. No.:              582           Length:        738
Score:                  37.00         Matches:       8
Percent Similarity:     72.73%        Conservative:  0
Best Local Similarity:  72.73%        Mismatches:    3
Query Match:            75.51%        Indels:        0
DB:                     10            Gaps:          0
```

US-09-823-649A-7 (1-11) x AV709998 (1-738)

```
Qy       1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
           |||   ||||||||||||||||      ||||||
Db     367 CTTTGCCAGAATTTGAACATTATCATGAAGGAA 399
```

Search completed: January 21, 2003, 12:24:30
Job time : 1172.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 21, 2003, 09:31:57 ; Search time 29.5714 Seconds
(without alignments)
165.854 Million cell updates/sec

Title: US-09-823-649A-7
Perfect score: 49
Sequence: 1 LAQNLNIXRKE 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09823649/runat_21012003_093151_24626/app_query.fasta_1.1393
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09823649_@CGN_1_1_77_@runat_21012003_093151_24626
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	39	79.6	2545	10	US-09-974-300-1524	Sequence 1524, Ap
	2	37	75.5	3177	10	US-09-801-574-3	Sequence 3, Appli
	3	36	73.5	181	10	US-09-960-352-13109	Sequence 13109, A
	4	36	73.5	2649	10	US-09-815-242-6571	Sequence 6571, Ap
	5	36	73.5	2787	10	US-09-815-242-6320	Sequence 6320, Ap
	6	36	73.5	8033	10	US-09-070-927A-121	Sequence 121, App
	7	35	71.4	151	10	US-09-878-574-8417	Sequence 8417, Ap
c	8	35	71.4	492	10	US-09-783-590-6288	Sequence 6288, Ap
	9	35	71.4	780	10	US-09-910-943-561	Sequence 561, App
c	10	34	69.4	239	10	US-09-867-701-8966	Sequence 8966, Ap
	11	34	69.4	684	10	US-09-974-300-6024	Sequence 6024, Ap
c	12	34	69.4	735	10	US-09-815-242-8266	Sequence 8266, Ap
c	13	34	69.4	753	10	US-09-815-242-4354	Sequence 4354, Ap
	14	34	69.4	1664	10	US-09-939-980-259	Sequence 259, App
	15	34	69.4	2670	10	US-09-815-242-9138	Sequence 9138, Ap
	16	34	69.4	4754	10	US-09-982-091A-1	Sequence 1, Appli
c	17	33	67.3	399	10	US-09-878-574-2036	Sequence 2036, Ap
	18	33	67.3	439	10	US-09-892-206-6	Sequence 6, Appli
	19	33	67.3	597	10	US-09-864-761-12301	Sequence 12301, A
c	20	33	67.3	1458	9	US-10-047-412A-3	Sequence 3, Appli
	21	33	67.3	1496	10	US-09-880-107-2424	Sequence 2424, Ap
c	22	33	67.3	1536	9	US-09-938-842A-222	Sequence 222, App
c	23	33	67.3	2170	9	US-10-047-412A-14	Sequence 14, Appl
c	24	33	67.3	2381	10	US-09-880-107-2271	Sequence 2271, Ap
c	25	33	67.3	2416	10	US-09-748-739A-3	Sequence 3, Appli
c	26	33	67.3	2416	10	US-09-748-739A-5	Sequence 5, Appli
c	27	33	67.3	2416	10	US-09-748-739A-7	Sequence 7, Appli
c	28	33	67.3	2416	10	US-09-748-739A-16	Sequence 16, Appl
	29	33	67.3	2657	10	US-09-892-206-1	Sequence 1, Appli
	30	33	67.3	50000	12	US-10-063-763-4	Sequence 4, Appli
c	31	32	65.3	446	10	US-09-867-701-7987	Sequence 7987, Ap
	32	32	65.3	695	10	US-09-764-877-716	Sequence 716, App
	33	32	65.3	768	9	US-09-938-842A-1461	Sequence 1461, Ap
c	34	32	65.3	995	12	US-10-021-509-4	Sequence 4, Appli
c	35	32	65.3	995	12	US-10-021-509-12	Sequence 12, Appl
	36	32	65.3	1278	10	US-09-770-445-22	Sequence 22, Appl
c	37	32	65.3	1730	9	US-10-028-072-7	Sequence 7, Appli
c	38	32	65.3	1802	10	US-09-908-711-132	Sequence 132, App
c	39	32	65.3	1996	10	US-09-887-576-40	Sequence 40, Appl
c	40	32	65.3	2000	9	US-09-938-842A-4751	Sequence 4751, Ap
c	41	32	65.3	2845	10	US-09-070-927A-462	Sequence 462, App
	42	32	65.3	3600	10	US-09-764-877-3729	Sequence 3729, Ap
	43	32	65.3	3600	10	US-09-764-877-3730	Sequence 3730, Ap
	44	32	65.3	8194	9	US-09-799-629-15	Sequence 15, Appl
	45	32	65.3	176373	9	US-10-095-407-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-974-300-1524
; Sequence 1524, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1524
; LENGTH: 2545
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1524

Alignment Scores:
Pred. No.: 14.6 Length: 2545
Score: 39.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1

```
Best Local Similarity:  72.73%          Mismatches:    2
Query Match:            79.59%          Indels:        0
DB:                     10              Gaps:          0

US-09-823-649A-7 (1-11) x US-09-974-300-1524 (1-2545)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          |||:::|||||||||   |||   |||||||||
Db   2170 TTGTCGCAAAACCTGGGAATTACAAGAAAGGAA 2202

RESULT 2
US-09-801-574-3
; Sequence 3, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Peijing Jeremy
;  APPLICANT:  Page, David C.
;  TITLE OF INVENTION: Reproduction-Specific Genes
;  FILE REFERENCE: 0399.2007-002
;  CURRENT APPLICATION NUMBER: US/09/801,574
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: 60/187,518
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: 60/261,557
;  PRIOR FILING DATE: 2001-01-12
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 3177
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-801-574-3

Alignment Scores:
Pred. No.:              53.9            Length:        3177
Score:                  37.00           Matches:       7
Percent Similarity:     81.82%          Conservative:  2
Best Local Similarity:  63.64%          Mismatches:    2
Query Match:            75.51%          Indels:        0
DB:                     10              Gaps:          0

US-09-823-649A-7 (1-11) x US-09-801-574-3 (1-3177)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          :::   |||||||||||||||   |||||||:::
Db   1880 ATAGCCAGAATCTGAATATAAAAACAAACAAA 1912

RESULT 3
US-09-960-352-13109
; Sequence 13109, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13109
;   LENGTH: 181
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 56-LIB3057-023-Q1-K1-F8
US-09-960-352-13109

Alignment Scores:
Pred. No.:              2.89            Length:        181
Score:                  36.00           Matches:       7
Percent Similarity:     90.00%          Conservative:  2
Best Local Similarity:  70.00%          Mismatches:    1
Query Match:            73.47%          Indels:        0
DB:                     10              Gaps:          0

US-09-823-649A-7 (1-11) x US-09-960-352-13109 (1-181)

Qy      2 AlaGlnAsnLeuAsnIle***ArgLysGlu 11
            |||:::|||||||||||   |||:::|||
Db     72 GCAAAGAATCTGAATATTACACGAGAGGAA 101

RESULT 4
US-09-815-242-6571
; Sequence 6571, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6571
;   LENGTH: 2649
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2649)
US-09-815-242-6571

Alignment Scores:
Pred. No.:              72.9            Length:        2649
Score:                  36.00           Matches:       7
Percent Similarity:     81.82%          Conservative:  2
Best Local Similarity:  63.64%          Mismatches:    2
Query Match:            73.47%          Indels:        0
DB:                     10              Gaps:          0

US-09-823-649A-7 (1-11) x US-09-815-242-6571 (1-2649)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          |||:::|||||||||   |||   |||||||:::
Db   2179 TTGTCCCAAAACTTAGGCATTACTAGAAAGCAA 2211

RESULT 5
US-09-815-242-6320
; Sequence 6320, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
```

```
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6320
;   LENGTH: 2787
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2787)
US-09-815-242-6320

Alignment Scores:
Pred. No.:                77.5          Length:          2787
Score:                    36.00         Matches:         8
Percent Similarity:       81.82%        Conservative:    1
Best Local Similarity:    72.73%        Mismatches:      2
Query Match:              73.47%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-7 (1-11) x US-09-815-242-6320 (1-2787)

Qy       1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
           ||||||:::  |||||||||  |||||||||
Db    2317 CTGGCGCGGCAATTGAACATTCCACGTAAAGAA 2349

RESULT 6
US-09-070-927A-121
; Sequence 121, Application US/09070927A
; Patent No. US20020120116A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles A. Kunsch
;                   Patrick J. Dillon
;                   Steven Barash
;        TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;        NUMBER OF SEQUENCES: 982
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/070,927A
;             FILING DATE: 04-May-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/046,655
;             FILING DATE: 1997-05-16
;             APPLICATION NUMBER: 60/044,031
;             FILING DATE: 1997-05-06
;             APPLICATION NUMBER: 60/066,009
;             FILING DATE: 1997-11-14
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB369
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 121:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8033 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 121:
US-09-070-927A-121

Alignment Scores:
Pred. No.:                277           Length:          8033
Score:                    36.00         Matches:         7
Percent Similarity:       81.82%        Conservative:    2
Best Local Similarity:    63.64%        Mismatches:      2
Query Match:              73.47%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-7 (1-11) x US-09-070-927A-121 (1-8033)

Qy       1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
           |||:::|||||||||  |||  |||||||:::
Db   7018 TTGTCTCAAAACTTAGGCATTACTACAAAGCAA 7050

RESULT 7
US-09-878-574-8417
; Sequence 8417, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 8417
;   LENGTH: 151
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701101161H1
US-09-878-574-8417

Alignment Scores:
Pred. No.:                3.91          Length:          151
Score:                    35.00         Matches:         7
Percent Similarity:       80.00%        Conservative:    1
Best Local Similarity:    70.00%        Mismatches:      2
Query Match:              71.43%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-7 (1-11) x US-09-878-574-8417 (1-151)
```

```
Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLys 10
          |||   |||||||||||||:::   ||||||
Db     93 TTGAAGCAGAACCTGAACTTGCACCGAAAA 122

RESULT 8
US-09-783-590-6288/c
; Sequence 6288, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Patrick J.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Li, Haodong
;  APPLICANT:  Rosen, Craig A.
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;  FILE REFERENCE: PO-16.2C1
;  CURRENT APPLICATION NUMBER: US/09/783,590
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 08/420,856
;  PRIOR FILING DATE: 1995-04-12
;  PRIOR APPLICATION NUMBER: 08/346,731
;  PRIOR FILING DATE: 1994-11-21
;  NUMBER OF SEQ ID NOS: 12485
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6288
;   LENGTH: 492
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (25)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (101)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (367)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (378)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (440)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (448)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (451)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (489)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-6288

Alignment Scores:
Pred. No.:                16.2          Length:          492
Score:                    35.00         Matches:         6
Percent Similarity:       88.89%        Conservative:    2
Best Local Similarity:    66.67%        Mismatches:      1
Query Match:              71.43%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-7 (1-11) x US-09-783-590-6288 (1-492)

Qy      3 GlnAsnLeuAsnIle***ArgLysGlu 11
          |||||||:::|||:::   |||||||||
Db    297 CAAAATATGAATCTTATTCGGAAAGAG 271

RESULT 9
US-09-910-943-561
; Sequence 561, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
;  APPLICANT: Hemmati-Brivanlou, Ali
;  APPLICANT:  Altman, Curtis
;  TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
;  FILE REFERENCE: 7529/1G148US1
;  CURRENT APPLICATION NUMBER: US/09/910,943
;  CURRENT FILING DATE:  2001-07-23
;  NUMBER OF SEQ ID NOS: 742
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 561
;   LENGTH: 780
;   TYPE: DNA
;   ORGANISM: Xenopus laevis
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(780)
;   OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-561

Alignment Scores:
Pred. No.:                28.2          Length:          780
Score:                    35.00         Matches:         7
Percent Similarity:       81.82%        Conservative:    2
Best Local Similarity:    63.64%        Mismatches:      2
Query Match:              71.43%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-7 (1-11) x US-09-910-943-561 (1-780)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          |||||||:::|||||||   |||   |||||||:::
Db    132 TTAGCGGAGAATTTAACTATCCATCGTAAACAA 164

RESULT 10
US-09-867-701-8966/c
; Sequence 8966, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aqlate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10212
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8966
;   LENGTH: 239
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-8966

Alignment Scores:
Pred. No.:                11.4          Length:          239
Score:                    34.00         Matches:         7
Percent Similarity:       72.73%        Conservative:    1
Best Local Similarity:    63.64%        Mismatches:      3
Query Match:              69.39%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-7 (1-11) x US-09-867-701-8966 (1-239)

Qy      1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
          |||||||:::|||||||         |||||||||
Db     78 CTTGCAGAGAACCTTGGCTTTCATAGAAAGGAA 46

RESULT 11
US-09-974-300-6024
; Sequence 6024, Application US/09974300
; Patent No. US20020146721A1
```

; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6024
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-6024

Alignment Scores:

Pred. No.:	40.5	Length:	684
Score:	34.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	69.39%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-7 (1-11) x US-09-974-300-6024 (1-684)

```
Qy       1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
           |||:::|||:::|||   |||   |||||||||
Db     218 CTCTCACAAAGCCTAGGCATTACGAGAAAAGAA 250
```

RESULT 12
US-09-815-242-8266/c
; Sequence 8266, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8266
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(735)
US-09-815-242-8266

Alignment Scores:

Pred. No.:	44.1	Length:	735
Score:	34.00	Matches:	6
Percent Similarity:	81.82%	Conservative:	3
Best Local Similarity:	54.55%	Mismatches:	2
Query Match:	69.39%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-7 (1-11) x US-09-815-242-8266 (1-735)

```
Qy       1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
           |||||||||||||||||||:::      ::::::
Db     307 CTCGCTCAAAATCTAAATTTAAATAATCGTGAT 275
```

RESULT 13
US-09-815-242-4354/c
; Sequence 4354, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4354
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4354

Alignment Scores:

Pred. No.:	45.4	Length:	753
Score:	34.00	Matches:	6
Percent Similarity:	81.82%	Conservative:	3
Best Local Similarity:	54.55%	Mismatches:	2
Query Match:	69.39%	Indels:	0
DB:	10	Gaps:	0

US-09-823-649A-7 (1-11) x US-09-815-242-4354 (1-753)

```
Qy       1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
           |||||||||||||||||||:::      ::::::
Db     340 CTCGCTCAAAATCTAAATTTAAATAATCGTGAT 308
```

RESULT 14

```
US-09-939-980-259
; Sequence 259, Application US/09939980
; Patent No. US20020082234A1
;   GENERAL INFORMATION:
;        APPLICANT: Black, Michael
;                   Burnham, Martin
;                   Hodgson, John
;                   Knowles, David
;                   Lonetto, Michael
;                   Nicholas, Richard
;                   Pratt, Julie
;                   Reichard, Richard
;                   Rosenberg, Martin
;                   Ward, Judith
;        TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynuclectides,
;                            Polypeptides and Their Uses
;        NUMBER OF SEQUENCES: 534
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SmithKline Beecham Corporation
;             STREET: 709 Swedeland Road
;             CITY: King of Prussia
;             STATE: PA
;             COUNTRY: USA
;             ZIP: 19406-0939
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/939,980
;             FILING DATE: 27-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/936,165
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Gimmi, Edward R
;             REGISTRATION NUMBER: 38,891
;             REFERENCE/DOCKET NUMBER: P50549
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 610-270-4478
;             TELEFAX: 610-270-5090
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 259:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1664 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: Genomic DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 259:
US-09-939-980-259

Alignment Scores:
Pred. No.:                118           Length:          1664
Score:                    34.00         Matches:         6
Percent Similarity:       81.82%        Conservative:    3
Best Local Similarity:    54.55%        Mismatches:      2
Query Match:              69.39%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-7 (1-11) x US-09-939-980-259 (1-1664)

Qy       1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
           ||||||||||||||||||:::      ::::::
Db     329 CTCGCTCAAAATCTAAATTTAAATAATCGTGAT 361

RESULT 15
US-09-815-242-9138
; Sequence 9138, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9138
;   LENGTH: 2670
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2670)
US-09-815-242-9138

Alignment Scores:
Pred. No.:                208           Length:          2670
Score:                    34.00         Matches:         7
Percent Similarity:       72.73%        Conservative:    1
Best Local Similarity:    63.64%        Mismatches:      3
Query Match:              69.39%        Indels:          0
DB:                       10            Gaps:            0

US-09-823-649A-7 (1-11) x US-09-815-242-9138 (1-2670)

Qy       1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
           |||:::   ||||||   |||   |||||||||
Db    2200 TTGTCTAATAATTTGGCAATTAGTCGTAAGGAA 2232


Search completed: January 21, 2003, 10:08:29
Job time : 31.5714 secs
```

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/54, 9/12, C12Q 1/68 — A1

(11) International Publication Number: **WO 98/40496**

(43) International Publication Date: 17 September 1998 (17.09.98)

(21) International Application Number: PCT/US98/05095

(22) International Filing Date: 12 March 1998 (12.03.98)

(30) Priority Data:
60/039,610 12 March 1997 (12.03.97) US

(71) Applicant: THE PERKIN–ELMER CORPORATION [US/US]; 850 Lincoln Centre Drive, Foster City, CA 94404 (US).

(72) Inventors: BRANDIS, John; 106 Sheffield, Hercules, CA 94547 (US). BLOOM, Curtis; 2631 Chalet Place, Chino Hills, CA 91709 (US). RICHARDS, Jack; 677 Deodar Lane, Bradbury, CA 91010 (US).

(74) Agent: BORTNER, Scott, D.; The Perkin–Elmer Corporation, 850 Lincoln Centre Drive, Foster City, CA 94404 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: DNA POLYMERASES HAVING IMPROVED LABELED NUCLEOTIDE INCORPORATION PROPERTIES

(57) Abstract

The present invention relates to mutant DNA polymerases that exhibit reduced discrimination against labeled nucleotides into polynucleotides. The DNA polymerases of the invention have at least one mutation in the nucleotide label interaction region of the enzyme such the mutation results in reduced discrimination against labeled nucleotides. The nucleotide label interaction regions is located at portions of the O–helix, (ii) the K–helix, and (iii) the inter O–P helical loop of Taq DNA polymerase or analogous positions in other DNA polymerases. In addition to providing novel mutant DNA polymerases, the invention also provides polynucleotides encoding the subject mutant DNA polymerases. The polynucleotides provided may comprise expression vectors for the recombinant production of the mutant polymerases. The invention also provides host cells containing the subject polynucleotides. The invention also includes numerous methods of using the subject DNA polymerases, including uses for chain termination sequencing and PCR. Another aspect of the invention is to provide kits for synthesizing fluorescently labeled polynucleotides in accordance with the methods of the invention. Kits of the invention comprise a mutant DNA polymerase of the invention and a fluorescently labeled nucleotide that exhibits reduced discrimination with respect to the mutant DNA polymerase in the kit.

E681

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

DNA POLYMERASES HAVING IMPROVED LABELED NUCLEOTIDE INCORPORATION PROPERTIES

Inventors: John Brandis, Curtis Bloom, and Jack Richards

Field of the Invention

The invention is related to DNA polymerases having mutations that alter the ability of the enzyme to incorporate labeled nucleotides into a polynucleotide molecule.

Background

DNA polymerases are enzymes that synthesize the formation of DNA molecules from deoxynucleotide triphosphates using a template DNA strand and a complementary synthesis primer annealed to a portion of the template. A detailed description of DNA polymerases and their enzymological characterization can be found in Kornberg, DNA Replication Second Edition, W. H. Freeman (1989).

DNA polymerases have a variety of uses in molecular biology techniques suitable for both research and clinical applications. Foremost among these techniques are DNA sequencing and nucleic acid amplification techniques such as PCR (polymerase chain reaction).

The amino acid sequence of many DNA polymerases have been determined. Sequence comparisons between different DNA polymerase have identified many regions of homology between the different enzymes. X-ray diffraction studies have determined the tertiary structures of Klenow fragment, T7 DNA polymerase, and Taq DNA polymerase. Studies of the tertiary structures of DNA polymerases and amino acid sequence comparisons have revealed numerous structural similarities between diverse DNA polymerases. In general, DNA polymerases have a large cleft that is thought to accommodate the binding of duplex DNA. This cleft is formed by two sets of helices, the first set is referred to as the "fingers" region and the second set of helices is referred to as the "thumb" region. The bottom of the cleft is formed by anti-parallel β sheets and is referred to as the "palm" region. Reviews of DNA polymerase structure can be found in Joyce and Steitz , Ann. Rev. Biochem. 63:777-822 (1994). Computer readable data files describing the three-dimensional structure of some

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

DNA polymerases have been publicly disseminated.

Fluorescently labeled nucleotides have greatly simplified and improved the utility of many procedures in molecular biology. The use of fluorescently labeled nucleotides for labeling polynucleotides in synthesis procedures, has to a large extent replaced the use of radioactive labeling. Fluorescently labeled nucleotides have been widely used in DNA sequencing, see Smith et al Nature 321:674-679 (1986), in PCR, and other forms of polynucleotide fragment analysis.

A major problem with using fluorescently labeled nucleotides is the ability of DNA polymerases to discriminate against the incorporation of fluorescently labeled nucleotides. For example, the inventors have discovered that in competition assays between a TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein) labeled 2' 3' dideoxynucleotide and the corresponding unlabeled dideoxynucleotide, Taq DNA polymerase incorporates the unlabeled dideoxynucleotide into DNA at least 85 times more frequently than the corresponding unlabeled nucleotide. This discrimination between labeled and unlabeled nucleotides has profound effects on procedures using DNA polymerases to label DNA. For example, much larger amounts of fluorescently labeled nucleotide must be used in sequencing reactions. This large amount of fluorescently labeled nucleotide is expensive and can generate excessive background fluorescence, thereby reducing the yield of sequence information.

In view of the problems arising from the ability of DNA polymerases to discriminate against the incorporation of fluorescently labeled nucleotides, the inventors have developed several novel DNA polymerases that have reduced discrimination against the incorporation of one or more fluorescently labeled nucleotides into DNA.

Summary

Naturally occurring DNA polymerases preferentially incorporate unlabeled nucleotides over corresponding fluorescently labeled nucleotides into polynucleotides. This ability of DNA polymerases to discriminate against fluorescently labeled nucleotide has undesirable effects on many molecular biology procedures that require the enzymatic addition of fluorescently labeled nucleotides, e.g., labeled dideoxy terminator sequencing. The present invention relates to mutant DNA polymerases that exhibit reduced discrimination against fluorescently labeled nucleotides into polynucleotides.

The DNA polymerases of the invention have at least one mutation in the nucleotide

label interaction region of the enzyme such that the mutation results in reduced discrimination against fluorescently labeled nucleotides. The nucleotide label interaction region of a DNA polymerase is formed by portions of the O-helix, (ii) the K helix, and (iii) the inter O-P helical loop of Taq DNA polymerase or analogous positions in other DNA polymerases. Amino acid residues within the nucleotide label interaction region as defined by TET (II) •ddC are E520. A531, L522, R523, E524, A525, H526, P527, I 528, V529, E530, K531, I532, R536, E537, R573, Q582, N583, V586, R587, P589, Q592, R593, R595, D610, T612, Q613, E615, R636, D637, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, H784, L817, E820, L828, K831, and E832. The sites at R660, T664, and E681 are of prefered sites for introducing mutations. In a preferred embodiment of the invention for use with fluorescein-type dyes, a mutation is present at position 681 converting an E (glutamic acid) to M (methionine), i.e., E681M. In a preferred embodiment of the invention for use with fluorescein- fluorescein energy transfer dyes a mutation is present at position 657 converting an L (leucine) to a G (glycine). In addition to providing mutant Taq DNA polymerases having reduced discrimination against labeled nucleotides, the invention includes mutants derived from a wide variety of DNA polymerases, both thermostable and otherwise.

In addition to providing novel mutant DNA polymerases, the invention also provides polynucleotides encoding the subject mutant DNA polymerases. The polynucleotides provided may comprise expression vectors for the recombinant production of the mutant polymerases. The invention also includes host cells containing the subject polymerase polynucleotides.

The invention also includes numerous methods of using the subject DNA polymerases. The subject methods involve synthesizing a fluorescently labeled polynucleotide by means of a polynucleotide synthesis reaction catalyzed by a mutant DNA polymerase that has reduced discrimination against incorporating labeled nucleotides into polynucleotides. The subject methods of polynucleotide synthesis include the step of extending a primed polynucleotide template with at least one fluorescent labeled nucleotide, wherein the extension is catalyzed by a DNA polymerase that has reduced discrimination against labeled nucleotides into polynucleotides. The subject methods of synthesizing a fluorescently labeled polynucleotide may be used in a variety of methods such as Sanger sequencing and the polymerase chain reaction (PCR).

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

Another aspect of the invention is to provide kits for synthesizing fluorescently labeled polynucleotides in accordance with the methods of the invention. Kits of the invention comprise a mutant DNA polymerase of the invention and a fluorescently labeled nucleotide that exhibits reduced discrimination with respect to the mutant DNA polymerase in the kit.

Brief Description of the Drawings

Figure 1 is a computer model of DNA bound to Taq DNA polymerase. Amino acid residues that form the nucleotide label interaction site are highlighted in orange. The rest of the polymerase is indicated in green. The template is indicated in blue. The dye moiety of the labeled nucleotide is red. The remaider of the labled nucleotide is white.

Figure 2 is plot of a next nucleotide effect assay.

Figure 3 is plot of a next nucleotide effect assay.

Figure 4 is a representation of the structure of the fluorescently labeled nucleotide "TET(II)•ddCTP."

Detailed Description of Specific Embodiments of the Invention.

Terminology

Positions of amino acid residues within a DNA polymerase are indicated by either numbers or number/letter combinations. The numbering starts at the amino terminus residue. The letter is the single letter amino acid code for the amino acid residue at the indicated position in the naturally occurring enzyme from which the mutant is derived. Unless specifically indicated otherwise, an amino acid residue position designation should be construed as referring to the analogous position in all DNA polymerases, even though the single letter amino acid code specifically relates to the amino acid residue at the indicated position in Taq DNA polymerase.

Individual substitution mutations are indicated by the form of a letter/number/letter combination. The letters are the single letter code for amino acid residues. The numbers indicate the amino acid residue position of the mutation site. The numbering system starts at the amino terminus residue. The numbering of the residues in Taq DNA polymerase is as described in U.S. Patent No. 5,079,352 (Gelfand). Amino acid sequence homology between different DNA polymerases permits corresponding positions to be assigned to amino acid residues for DNA polymerases other than Taq. Unless indicated otherwise, a given number

refers to position in Taq DNA polymerase. The first letter, i.e., the letter to the left of the number, represents the amino acid residue at the indicated position in the non-mutant enzyme. The second letter represents the amino acid residue at the same position in the mutant enzyme. For example, the term "R660D" indicates that the arginine at position 660 has been replaced by an aspartic acid residue.

The term "discrimination" as used herein refers to the property of a DNA polymerase to preferentially incorporate unlabeled nucleotides over corresponding fluorescently labeled nucleotides into DNA, i.e., the DNA polymerase discriminates against the fluorescently labeled nucleotide. Preferential incorporation may be measured in an assay in which a fluorescently labeled 2'3' dideoxynucleotide and a corresponding unlabeled 2'3' dideoxynucleotide compete for incorporation into the same site of a polynucleotide. An example of such an assay can be found below in example 2.

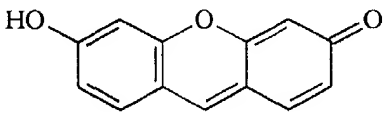
The term "reduced discrimination" as used herein refers to reduction in discrimination against incorporation of a fluorescently labeled nucleotides in a mutant DNA polymerase as compared to the parent enzyme. A reduction in discrimination may be described quantitatively by reference to the selectivity assays in Example 2 or reference to other assays providing for measurement of the same properties of the polymerase. A reduction in selectivity number as measured by the selectivity assays is a reduction in discrimination and may be expressed by a ratio of selectivity numbers. For example, a mutant DNA polymerase with a selectivity number of 8 would have a 10-fold reduction in discrimination when compared with a parent DNA polymerase having a selectivity number of 80.

The term "parent" or "parent enzyme" is used to distinguish a mutant DNA polymerase from the DNA polymerase that the mutant enzyme was derived from. Thus any naturally occurring DNA polymerase may be referred to as parent enzyme. A first DNA polymerase having mutations with respect to a naturally occurring enzyme is also be referred to as a parent enzyme with respect to a second DNA polymerase having additional mutations.

The term "discrimination reducing mutations" refers to mutations in the nucleotide label interaction region of a DNA polymerase that result in reduced discrimination against the incorporation of fluorescently labeled nucleotides. The term is used to distinguish mutations in a DNA polymerase, including mutations in the nucleotide label interaction region, that do not reduce discrimination against fluorescently labeled nucleotides from mutations that do reduce discrimination.

The term "nucleotide" as used herein, unless specifically noted otherwise, is used broadly to refer to both naturally occurring nucleotide and a variety of analogs including 2',3' dideoxynucleotides.

The term "fluorescein-type dyes" refers to a class of xanthene dye molecules which include the following fused three-ring system:



where a wide variety of substitutions are possible at each deoxy ring position. A particularly preferred subset of fluorescein-type dyes include the 4,7,-dichorofluoresceins (Menchen). Examples of fluorescein-type dyes used as fluorescent labels in DNA sequencing methods include 6-carboxyfluorescein (6-FAM), 5-carboxyfluorescein (5-FAM), 6 -carboxy-4,7,2',7'-tetrachlorofluorescein (TET), 6-carboxy-4,7,2',4',5',7'-hexachlorofluorescein (HEX), 5-(and 6)carboxy-4',5'-dichloro-2'7'-dimethoxyfluorescein (JOE), and 5-carboxy-2',4',5',7'-tetrachlorofluorescein (ZOE). Many times the designation -1 or -2 is placed after an abbreviation of a particular dye, e.g., HEX-1. The "-1" and "-2" (or "I" and "II") designations indicate the particular dye isomer being used. The 1 and 2 isomers are defined by the elution order (the 1 isomer being the first to elute) of free dye in a reverse-phase chromatographic separation system utilizing a C-8 column and an elution gradient of 15% acetonitrile/85% 0.1 M triethylammonium acetate to 35% acetonitrile / 65% 0.1 M triethylammonium acetate.

The term "alkynylamino type linker" refers to an alkynylamino linker of the type as described in U.S. Patent No. 5,047,519 (Hobbs), U.S. Patent No. 5,151,507 (Hobbs), and U.S. Patent Application No. 08/696,808, filed August 13, 1996. Additional alkynylamino type liners are described in U. S. Patent Application No. 08/833,855, filed April 10, 1997.

The term "TET(II)•ddCTP" refers to the fluorescently labeled nucleotide of the structure indicated in figure 4.

The term "fluorescence energy transfer dye" refers to dye moieties joined by a linker that permits fluorescence energy transfer between the two dye moieties. For use in chain termination sequencing, the linker is sufficiently small and of the proper shape and orientation

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

to permit a DNA polymerase to incorporate a nucleotide triphosphate labeled with the dye interest. Examples of energy transfer dyes can be found in European Patent Application No. EP 0 805 140, U.S. Patent Application No. 08/642,330 (filed May 3, 1996), and U.S. Patent Application No. 08/726,462 (filed October 4, 1996).

The term "mutation" as used herein refers to a change in amino acid residue at a specific location of a protein. The change in amino acid residue is a change defined with respect to a naturally occurring protein. A protein having a mutation may be referred to as a "mutant" protein.

Embodiments of the Invention

The present invention relates to DNA polymerases containing mutations that reduce the ability of the polymerase to discriminate against the incorporation of fluorescently labeled nucleotides into polynucleotides. These mutations are in a region of the DNA polymerase molecule referred to herein as "the nucleotide label interaction region." The nucleotide label interaction region is formed by portions of three regions of the DNA polymerase. These three regions are located in (i) the O-helix, (ii) the K helix, and (iii) the inter O-P helical loop of Taq DNA polymerase or analogous positions in other DNA polymerases. DNA polymerases having reduced discrimination against fluorescently labeled nucleotides are particularly useful for chain termination DNA sequencing using 2'3' dideoxynucleotides, i.e., Sanger type sequencing.

Enzyme kinetic experiments (described in examples 2 and 3) performed with Taq DNA polymerase and fluorescently labeled dideoxynucleotides support a theory that Taq DNA polymerase and other DNA polymerases, undergo a conformational shift upon the binding of nucleotides during DNA synthesis. This predicted conformational shift suggests a set of amino acid residues that interact with fluorescent labels joined by a linker to the nucleic acid base of a nucleotide, thereby resulting in discrimination against nucleotides that are fluorescently labeled. This set of amino acid residues forms the nucleotide label interaction region. The specific molecular model for the binding of fluorescently labeled nucleotide to a DNA polymerase proposed by the applicants is used to predict the amino acid residues that form the nucleotide label interaction region of a given DNA polymerase Applicants model for a conformational shift in DNA polymerase during DNA synthesis is offered as a explanation of how the nucleotide label interaction region was determined. The

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

model provides guidance in making mutations in DNA polymerase that reduce the ability of a DNA polymerase to discriminate against the incorporation of fluorescently labeled nucleotides into polynucleotides. Figure 1 is a computer model showing how DNA and Taq DNA polymerase interact in the model. Whether or not the true mechanism of DNA polymerase-nucleotide interaction is the same or different as the model used to determine the parameters of the nucleotide label interaction region is not determinative to the operability of the invention described herein.

The mutant DNA polymerases of the invention exhibit reduced discrimination against nucleotides labeled with a fluorescein-type dye. In other words, the mutant DNA polymerases of the invention contain at least one mutation that increases the ability of the polymerase to incorporate a fluorescein-type dye labeled nucleotide relative to the corresponding unlabeled nucleotide. In addition to reduced discrimination against nucleotides labeled with fluorescein-type dyes, the mutant DNA polymerases of the invention may also exhibit reduced discrimination against nucleotides labeled with other fluorescent dyes that are not fluorescein-type dyes, as well as reduced discrimination against other detectable moieties. The fluorescently labeled nucleotides for which a given embodiment of the mutant DNA polymerases of the invention exhibit reduced discrimination may vary with respect to the particular fluorescent label, the linker used to attach the fluorescent label to the nucleotide, the site of attachment for the linker on the fluorescent label, the specific nucleotide base that is selected, and the site of attachment for the linker on the nucleotide. The precise degree of reduction in discrimination against a fluorescently labeled nucleotide will vary in accordance with the specific mutation or mutations introduced into the DNA polymerase. The precise degree of reduction in discrimination will also vary in accordance with the specific fluorescently labeled nucleotide assayed, e.g., variations in base, dye, or linker. Mutant DNA polymerase of the invention may exhibit anywhere from a slight reduction in discrimination against fluorescently labeled nucleotides to a complete elimination in discrimination, i.e., the mutant enzyme does not significantly differ with respect of rate of incorporation of labeled or unlabeled nucleotides. It is preferable to use embodiments of the subject mutant DNA polymerases that have at least a two-fold reduction in discrimination against one or more fluorescein type dye labeled nucleotides.

It will be appreciated by persons skilled in the art of molecular biology that the nucleotide label interaction region of a given DNA polymerase is defined with respect to a

specific fluorescently labeled nucleotide. Changes in one or more of the following parameters of the structure of a fluorescently labeled nucleotide may alter the identity of the amino acid residues that form the nucleotide label interaction site of a given DNA polymerase: (1) identity of the base, (2) the site of attachment on the nucleotide base, (3) the identity of the linker joining the base to the florescent dye, and (4) the identity of the fluorescent dye. The nucleotide labeled interaction region of Taq defined with respect to TET(II)•ddCTP comprises the amino acid residues E520. A531, L522, R523, E524, A525, H526, P527, I 528, V529, E530, K531, I532, R536, E537, R573, Q582, N583, V586, R587, P589, Q592, R593, R595, D610, T612, Q613, E615, R636, D637, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, H784, L817, E820, L828, K831, and E832. The sites at R660, T664, and E681 are of prefered sites for introducing mutations. Given that the 3-dimensional structure of Taq DNA polymerase (and other DNA polymerases) is well known and the three dimensional structure of TET(II)•ddCTP is understood with a high degree of certainty, the location of the amino acid residues that constitute the labeled nucleotide interaction region with respect to TET(II)•ddCTP may be translated to a different set of amino acid residues to accommodate structural differences between TET(II)•ddCTP and other fluorescently labeled nucleotides so as to define the labeled nucleotide interaction site with respect to those other nucleotides. For example, increasing the length of the linker between the base and the fluorescent label and the base may predictably alter the identity of amino acid residues that form the labeled nucleotide interaction site, even though the base, base attachment site, and fluorescent dye are the same. In many embodiments of the subject polymerases, the set of amino acid residues that form the labeled nucleotide interaction site with respect to a given fluorescently labeled nucleotide will overlap with the set of amino acid residues that form the labeled nucleotide interaction site as defined with respect to a second fluorescently labeled nucleotide.

Embodiments of the invention include mutant DNA polymerases that exhibit reduced discrimination against nucleotides labeled with fluorescein-type dyes, wherein the fluorescein type dye is joined to the nucleotide base by an alkynylamino-type linker. The fluorescein-type dye may be a fluorescent energy transfer dye, comprising a fluorescein-type dye moiety as a component of the energy transfer dye. In addition to reduced discrimination against fluorescently labeled nucleotides comprising an alkynylamino-type linker, the mutant DNA polymerases of the invention may also exhibit reduced discrimination against nucleotides

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

comprising other types of linker. In order to minimize stearic interference between the polynucleotide and the fluorescent label, purines are usually labeled at position 7 and pyrimidines are usually labeled at position 5.

Mutant DNA polymerases of the invention have one or more discrimination reducing mutations at amino acid residue positions within the nucleotide label interaction region of a given DNA polymerase. Discrimination reducing mutations are usually, although not necessarily, substitution mutations. Several different amino residues may be substituted at a given position of a parent enzymes so as to give rise to a discrimination reducing mutations. The amino acid residues at a given residue position within the nucleotide label interaction region may be systematically varied so as to determine which amino acid substitutions result in the reduction of discrimination against the fluorescein-type dye labeled nucleotide dye of interest and the degree of such a reduction in discrimination. The extent to which a particular mutation (or set of mutations) reduces discrimination may be measured by a selectivity assay as described in example 2. The substitution mutation is preferably, although not necessarily, a mutation that reduces the size of the amino acid residue side chain of the amino acid residue present in the parent DNA polymerase. Mutations are preferably, although not necessarily, conservative so as to maintain the specific polar or non-polar character of the amino acid residue at the analogous position parent molecule. The mutations in the nucleotide label interaction region of a DNA polymerase preferably result in the substitution of the amino acid residue of the parent enzyme with the amino acid residue at the corresponding position of phage T7 DNA polymerase (provided that a difference exists between the amino acid residues at that position in T7 polymerase and the parent enzyme).

Discrimination reducing mutations are in the nucleotide label interaction region of DNA polymerases. The nucleotide label interaction region is formed by portions of three regions of the DNA polymerase. These three regions are located in (i)the O-helix, (ii) the K helix, and (iii) the inter O-P helical loop of Taq DNA polymerase or analogous positions in other DNA polymerases. Positions in Taq DNA polymerase that form the nucleotide label interaction region are positions E520. A531, L522, R523, E524, A525, H526, P527, I 528, V529, E530, K531, I532, R536, E537, R573, Q582, N583, V586, R587, P589, Q592, R593, R595, D610, T612, Q613, E615, R636, D637, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, H784, L817, E820, L828, K831, and E832. Analogous positions in DNA polymerases other than Taq are also form

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

a nucleotide label interaction region. Preferred positions for substitution mutations are R595, D655, R660, and E681. A particularly preferred position for mutations is E681, with the preferred substitution at position 681 being M. Other suitable substitution mutations at E681 are as follows (listed in order of decreasing preference, excpt where note by a equal sign to denote approximate equivalence"): M>I>W>L>V>P>H=K=G=T=S>D=A=N>Y=C. A preferred substitution mutation at position R660 is R660D.

The specific amino acid residues that form the nucleotide interaction region will vary in accordance with the particular DNA polymerase selected as a parent enzyme for the introduction of discrimination reducing mutations. The determination of analogous amino acid residues positions between different DNA polymerases may easily be achieved by the person skilled in the art because of the large number of DNA polymerase amino acid sequences that have been determined and the many regions of homology have been found between these different DNA polymerases. For example, a large compilation of the amino acid sequences of DNA polymerases from a wide range of organism and homology alignments between the sequences can be found in Braithwaite and Ito, Nucl. Acids Res. 21(4):787-802 (1993). Examples of amino acid residues within the nucleotide label interaction regions of phage T7 polymerase and *E. coli* DNA polymerase are provided in Table 1. In addition to providing mutant DNA polymerases having reduced discrimination for fluorescein type dyes in Taq, T7 and *E. coli* DNA polymerase I, the invention provides mutant DNA polymerases from many other organisms. In general, the teachings of the invention may used to produce mutant DNA polymerases having reduced discrimination for fluorescein type dyes from any DNA polymerase that shares sufficient amino acid sequence homology to Taq DNA polymerase to permit a person of ordinary skill in the art to identify one or more amino acid residue positions in the DNA polymerase that are analogous to positions E520. A531, L522, R523, E524, A525, H526, P527, I 528, V529, E530, K531, I532, E537, R573, V586, R587, P589, Q592, R593, R595, D610, T612, Q613, E615, R636, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, L817, E820, L828, K831, and E832 in Taq DNA polymerase. Parent DNA polymerases that may be modified to contain discrimination reducing mutations in the nucleotide label interaction region include, but are not limited to, DNA polymerases from organisms such as *Thermus flavus*, *Pyrococcus furiosus*, *Thermotoga neapolitana*, *Thermococcus litoralis*, *Sulfolobus solfataricus*, *Thermatoga maritima*, *E. coli* phage T5, and *E. coli* phage T. The DNA

polymerases may be thermostable or not thermostable. It will be appreciated that the present invention enables persons skilled in the art to introduce fluorescein-type dye discrimination reducing mutations in to DNA polymerases from a wide variety of organisms, including DNA polymerases that have not been isolated at the time of the filing of this application provided. Additionally, embodiments of the invention includes some purified naturally-occurring DNA polymerases that have the desired low degree of discrimination against fluorescently labeled nucleotides. Such naturally-occurring DNA polymerases are structurally and functionally analogous to the mutant DNA polymerases explicitly described herein.

The amino acid residues that constitute the nucleotide label interaction region of a given DNA polymerase vary in accordance with the specific fluorescently labeled nucleotide that is used to define the nucleotide label interaction region. Similarly, the mutations that are discrimination reducing mutations may vary in accordance with the specific fluorescently labeled nucleotide that is used to define the labeled nucleotide interaction region. Additionally, the degree of discrimination reduction achieved by the mutation (or mutations) in the labeled nucleotide interaction site may vary with the specific labeled nucleotide of interest. For example, E681M is the preferred discrimination reducing mutation in Taq with respect to TET(II)•ddCTP resulting in a 47x reduction in discrimination and a significantly lower reduction in discrimination against a second fluorescently labeled nucleotide. Conversely, an E681T mutation may result in a high level reduction in discrimination against the second fluorescently labeled nucleotide and only a low level of reduction in discrimination against TET(II)•ddCTP.

Given that a mutant DNA polymerase of the invention may have discrimination reducing mutation in the nucleotide label interaction region resulting in a significant degree of reduction in discrimination for a specific fluorescently labeled nucleotide and little or no reduction in the degree of reduction of discrimination against another fluorescently labeled nucleotide (assuming there is significant discrimination against that fluorescently labeled nucleotide by the parent DNA polymerase), a given mutant DNA polymerase may be said to be "receptive" with respect to one or more given fluorescently labeled nucleotide. A specific mutant DNA polymerase is referred to as "receptive" with respect to a specific fluorescently labeled nucleotide if a discrimination reducing mutation in the nucleotide label interaction site in the specific enzyme of interest results in at least a five fold reduction in discrimination against that given fluorescently labeled nucleotide. A mutant DNA polymerase of the

invention may be receptive with respect to more than one fluorescently labeled nucleotide. Conversely, a specific fluorescently labeled nucleotide may be "receptive" with respect to a given mutant DNA polymerase of the invention.

In embodiments of the subject mutant DNA polymerases comprising more than one discrimination reducing mutation in the nucleotide label interaction region, the mutation site may be in the same or different region of the three regions of a polymerase that form the nucleotide label interaction region. In general, mutant DNA polymerases of the invention will have 1, 2, or 3 discrimination reducing mutations. However, the invention also provides mutant DNA polymerases having more than 3 discrimination reducing mutations. By combining multiple discrimination reducing mutations, greater levels of reduction in labeled nucleotide discrimination may be achieved. However, in many embodiments of the invention, mutant DNA polymerases have levels of reduced labeled nucleotide discrimination that are the same or less than the levels of DNA polymerase with single discrimination reduction mutations in the nucleotide label interaction region. Preferred combinations of mutations in a Taq DNA polymerase background are R660D, E681G, and F667Y, i.e., Taq DNA polymerase mutant (R660D, E681G, and F667Y).

Different embodiments of DNA polymerase having mutations in the nucleotide label interaction region differ with respect to the degree of reduction in discrimination against specific fluorescently labeled nucleotides. These differences may be measured by an assay in order to determine which specific embodiments have the greatest degree of reduction in discrimination against the particular fluorescently labeled nucleotides of interest. Generally, such assays measure competition between a fluorescently labeled nucleotide and an unlabeled nucleotide for incorporation into the same site on a primed template. One example of such an assay (referred to herein as a "selectivity assay") is described in detail below in Example 2.

The mutant DNA polymerases of the invention may comprise numerous mutations in addition to discrimination reduction mutations in the nucleotide label interaction region. These secondary mutations may be either inside or outside the nucleotide label interaction region. Secondary mutations may be selected so as to have as to confer some useful property on the mutant DNA polymerase. For example, additional mutations may be introduced to increase thermostability; decrease thermostability, increase processivity, decrease processivity, decrease 3'-5' exonuclease activity, increase 3'-5' exonuclease activity, decrease 5'-3'

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

exonuclease activity, increase 5'-3' exonuclease activity, and increase incorporation of dideoxynucleotides. Alternatively, the secondary mutations may be essentially neutral in known effect.

Of particular interest are embodiments of the subject mutant DNA polymerase that comprise one or more secondary mutation that reduce 3'-5' exonuclease activity. DNA polymerases that are deficient in 3'-5' exonuclease activity have superior properties for PCR and for chain termination polynucleotide sequencing. Mutations that reduce 3'-5' exonuclease activity in DNA polymerase are well known to person of ordinary skill in the art. Detailed guidance on how to introduce mutations that reduce 3'-5' exonuclease activity can be found, among other places in U.S. Patent No. 4,795,699 (Tabor); U.S. Patent No. 5,541,099; U.S. Patent No. 5,489,523; and Bernad et al., Cell 59:219-288 (1989). Examples of such mutations in Taq DNA polymerase include G46D. For embodiments of the mutant DNA polymerases that are used for sequencing, it is preferable to include a G46D (or analogous mutations in DNA polymerases other than Taq) in addition to mutations in the nucleotide label interaction region.

Also of interest among secondary mutations in the subject DNA polymerase mutants are mutations that increase incorporation of dideoxynucleotides, i.e., reduce the ability of a DNA polymerase to discriminate against dideoxynucleotide as opposed to deoxynucleotides. Guidance on making such mutations can be found, among other places in published PCT application WO96/12042 (application number PCT/US95/12928). Of particular interest is the mutation F667Y in Taq and analogous mutations in other DNA polymerase. While F667Y is not part of the nucleotide label interaction region in Taq DNA polymerase with respect to Tet(II)•ddLTP, F667Y mutations may reduce discrimination against fluorescein-type dye labeled nucleotides (see Table 1). Accordingly, for use in certain procedures, e.g., DNA sequencing, be desirable to combine an F667Y mutations with one or more discrimination reducing mutations in the nucleotide label interaction region so as to reduce discrimination of the polymerase between deoxynucleotides and 2'3' dideoxynucleotides. Mutant DNA polymerase of the invention having the F667Y mutation (or equivalent thereof) are particularly useful in Sanger type DNA sequencing with fluorescently labeled 2'3' dideoxynucleotide chain terminators.

Numerous genes encoding DNA polymerases have been isolated and sequenced. This sequence information is available on publicly accessible DNA sequence databases such as

GENBANK. A large compilation of the amino acid sequences of DNA polymerases from a wide range of organism can be found in Braithwaite and Ito, Nucl. Acids Res. 21(4):787-802 (1993). This information may be used in designing various embodiments of DNA polymerases of the invention and polynucleotide encoding these enzymes. The publicly available sequence information may also be used to clone genes encoding DNA polymerases through techniques such as genetic library screening with hybridization probes.

Other embodiments of the invention are polynucleotide sequences encoding the mutant DNA polymerases provided herein. Polynucleotide sequences encoding the mutant DNA polymerase of the invention may be used for the recombinant production of the mutant DNA polymerases. Polynucleotide sequences encoding mutant DNA polymerases having reduced discrimination against fluorescently labeled nucleotide may be produced by a variety of methods. A preferred method of producing polynucleotide sequences encoding mutant DNA polymerases having reduced discrimination against fluorescently labeled nucleotides is by using site-directed mutagenesis to introduce desired discrimination reducing mutations into polynucleotides encoding the parent DNA polymerase molecules. Site-directed mutagenesis techniques are well known in the art as exemplified by U.S. Patent No. 4,711,848; U.S. Patent No. 4,873,192; U.S. Patent No. 5,071,743; U.S. patent, 5,284,760; U.S. Patent No. 5,354,670; U.S. Patent No. 5,556,747; Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982), and Edelman et al DNA 2:183 (1983). Detailed protocols for site-directed mutagenesis are also given many general molecular biology textbooks such as Sambrook et al Molecular Cloning a Laboratory Manual 2nd Ed. Cold Spring Harbor Press, Cold Spring Harbor (1989), Ausubel et al. Current Protocols in Molecular Biology, (current edition). Additionally, many text books on PCR (the polymerase chain reaction), such as Diefenbach and Dveksler, PCR Primer: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY (1995), describe methods of using PCR to introduce directed mutations. Genes encoding parent DNA polymerase may be isolated using conventional cloning techniques in conjunction with publicly-available sequence information. Alternatively, many cloned polynucleotide sequences encoding DNA polymerases have been deposited with publicly-accessible collection sites, e.g., the American type culture collection deposit accession number ATCC 40336 is a phage clone of Taq DNA polymerase.

In addition to producing the mutant DNA polymerase encoding polynucleotides of the invention by introducing directed mutations into polynucleotides encoding parent DNA

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

polymerases, it is possible (although difficult) to produce the polynucleotides of the invention primarily by *in vitro* DNA synthesis techniques. *In vitro* DNA synthesis techniques are well known to those skilled in the art and examples of *in vitro* DNA synthesis can be found in U.S. Patent No. 5,252,530; U.S. Patent No. 4,973,679; U.S. Patent No. 5,153,319; U.S. Patent No. 4,668,777; U.S. Patent No. 4,500,707; U.S. Patent No. 5,132,418; U.S. Patent No. 4,415,732; U.S. Patent No. 4,458,066; and U.S. Patent No. 4,811,218. When producing relative polynucleotide molecules by *in vitro* DNA synthesis, smaller molecules are usually produced first and subsequently joined together by hybridization and ligation. Mutant DNA polymerase encoding polynucleotides may also be produced by a combination of *in vitro* synthesis and site-directed mutagenesis of cloned genes.

Polynucleotide encoding the mutant DNA polymerase of the invention may be used for the recombinant expression of the mutant DNA polymerases. Generally, the recombinant expression of the mutant DNA polymerase is effected by introducing a mutant DNA polymerase into an expression vector adapted for use in particular type of host cell. Thus, another aspect of the invention is to provide expression vectors comprising a polynucleotide encoding a mutant DNA polymerase of the invention, such that the polymerase encoding polynucleotide is functionally inserted int the expression vector. The invention also provide host cells comprising the expression vectors of the invention. Host cells for recombinant expression may be prokaryotic or eukaryotic. Example of host cells include bacterial cells, yeast cells, cultured insect cell lines, and cultured mammalian cells lines. Preferably, the recombinant host cell system is selected so as to closely match the organism from which the mutant DNA polymerase was derived. For example, prokaryotic DNA polymerases are preferably expressed in a prokaryotic expression system. A wide range of expression vectors are well known in the art. Description of various expression vectors and how to use them can be found among other places in U.S. Patent No. 5604118; U.S. 5,583,023; U.S. Patent No. 5,432,082; U.S. Patent No. 5,266,490; U.S. Patent No. 5,063,158; U.S. Patent No. 4,966,841; U.S. Patent No. 4,806,472; U.S. Patent No. 4,801,537; and Goedel et al., Gene Expression Technology, Methods of Enzymology, Vol. 185, Academic Press, San Diego (1989). The expression of DNA polymerases in recombinant cell systems is a well-established technique. Examples of the recombinant expression of DNA polymerase can be found in U.S. Patent No. 5,602,756; U.S. Patent No. 5,545,552; U.S. Patent No. 5,541,311; U.S. Statutory Inventor Registration H1,531; U.S. Patent No. 5,500,363; U.S. Patent No. 5,489,523; U.S.

Patent No. 5,455,170; U.S. Patent No. 5,352,778; U.S. Patent No. 5,322,785; and U.S. Patent No. 4,935,361.

Other embodiments of the invention include multiple DNA polymerase compositions particularly useful for polynucleotide sequencing, such compositions comprise at least two different mutant DNA polymerases of the invention, wherein (1) the first mutant DNA polymerase is receptive with respect to a first fluorescently labeled nucleotide; (2) the second mutant DNA polymerase is receptive with respect to a second fluorescently labeled nucleotide; and (3) the first and second fluorescently labeled nucleotides differ from one another with respect to their nucleotide bases and fluorescent labels. The first and second fluorescently labeled bases may also differ with respect to one another by way of the linker, the base attachment position, or the fluorescent dye attachment site. The subject compositions are useful for catalyzing the sequencing reactions in Sanger type DNA sequencing with fluorescent dye labeled 2'3' dideoxy chain terminating nucleotides. Chain termination sequencing with fluorescently labeled terminators preferably employs at least two, and more preferably 4 different fluorescently labeled chain terminators, wherein each different base is labeled with a distinctive fluorescent label. Because of the necessary structural differences between the different fluorescently labeled chain terminators required for a sequencing reactions, i.e., nucleotide bases and fluorescent labels, there are many mutant DNA polymerases of the invention that are not receptive to all of the fluorescently labeled terminators necessary for a given sequencing reaction. Thus, there are embodiments of the subject DNA polymerases that may have undesirably high levels of discrimination against one or more of the labeled terminators used in a sequencing reaction set. The subject compositions of two or more mutant polymerases ameliorates this problem by simultaneously employing multiple mutant DNA polymerases that are receptive to different chain labeled terminators, thereby having at least one of the mutant polymerases "compensate" for the discrimination against a particular fluorescently labeled terminator by the other polymerases catalyzing the sequencing reactions. The ratio of the different DNA polymerases in the composition preferably are selected so as to result in approximately equal levels of total activity for each of the different mutant DNA polymerases. Differences in specific activity between the different mutant polymerases may be taken into account when equalizing total activity ratios between the polymerases. Differences in activity levels between the various mutant DNA polymerases in the subject compositions may also be compensated for by

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

adjusting the levels of the different fluorescently labeled terminators in the subject compositions. The subject multiple polymerase compositions may comprise two, three, four, or more different mutant DNA polymerases. The mutant polymerase may or may not be derived from the same species or strain. The different mutation DNA polymerases in the subject mutant polymerase compositions may or may not be receptive for one or more of the fluorescently labeled nucleotides in a given set fluorescently labeled dideoxynucleotides for sequencing.

The invention also includes various methods of using the mutant DNA polymerases (or subject multiple mutant DNA polymerase compositions) of the invention. The mutant DNA polymerases of the invention may be substituted for the corresponding parent DNA polymerases in most procedures that employ DNA polymerases. In order to more fully take advantage of the properties of the subject mutant DNA polymerases, the amount (or concentration) of labeled and unlabeled nucleotides used in the methods of the invention may be changed with respect to the amounts (or concentrations) used in the corresponding methods employing convention DNA polymerases. These changes in the amount of nucleotide may be optimized by routine experimentation. Methods of the invention comprise the step of extending a primed polynucleotide template with at least one fluorescently labeled nucleotide, wherein the extension is catalyzed by a mutant DNA polymerase of the invention. Thus, the subject methods result in the formation of one or more different fluorescently labeled polynucleotides produced by primer extension. The subject methods of synthesizing a fluorescently labeled polynucleotide may be used in a variety of procedures including, but not limited to, Sanger sequencing (e.g., dideoxy nucleotide chain termination), the polymerase chain reaction (PCR), polynucleotide labeling, minisequencing. The reduced discrimination against fluorescently labeled nucleotide properties of the subject mutant DNA polymerase is particularly useful for Sanger DNA sequencing reactions, including cycle sequencing. The use of the subject mutant DNA polymerases for Sanger sequencing reduces the amount of fluorescently labeled chain terminating nucleotides required for a sequencing reaction an may in many case be used to increase the number of bases that may be identified in single sequencing reaction that is analyzed on an automated fluorescence-based sequencing apparatus such as an Applied Biosystems 310 or 377(Applied Biosystems Division of Perkin-Elmer, Foster City, CA.). Detailed protocols for Sanger sequencing are known to those skilled in the art and may be found, for example in Sambrook et al, Molecular Cloning, A

Laboratory Manual, Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989).

The invention also provides kits for synthesizing fluorescently labeled polynucleotides. The kits may be adapted for performing specific polynucleotide synthesis procedures such as DNA sequencing or PCR. Kits of the invention comprise a mutant DNA polymerase of the invention and a fluorescently labeled nucleotide that exhibits reduced discrimination with respect to the mutant DNA polymerase in the kit. Kits preferably contain detailed instructions on how to perform the procedures for which the kits are adapted. Optionally, the subject kit may further comprise at least one other reagent required for performing the method the kit is adapted to perform. Examples of such additional reagents include unlabeled nucleotides, buffers, cloning vectors, restriction endonucleases, sequencing primers, and amplification primers. The reagents include in the kits of the invention may be supplied in premeasured units so as to provide for greater precision and accuracy.

Other embodiments of the invention include kits comprising (1) the subject compositions of multiple mutant DNA polymerases, and (2) fluorescently labeled chain terminating nucleotides suitable for use with the subject compositions, i.e., each labeled chain terminator is receptive with respect to at least one of the mutant DNA polymerases in the composition. Additional embodiments of the invention include kits for sequencing DNA that comprise a multiple mutant polymerase composition of the invention and at least two different fluorescently labeled chain terminating nucleotides are labeled at different bases, wherein each of the fluorescently labeled chain terminating nucleotides is receptive with respect to at least one mutant DNA polymerase in the composition.

The invention, having been described above, may be better understood by reference to the following examples. The examples are offered, for among other reasons, to illustrate specific embodiment of the invention and should not be construed as a limitation on the invention.

EXAMPLES

Example 1

Purification of Mutant Forms of Taq DNA Polymerase

Lysates of *E. coli* containing recombinant constructs designed for the productiion of

recombinant mutant Taq DNA polymerases were made essentially as described in tDesai, U.J. and Pfaffle, P.K., Biotechniques, 19:780-784 (1995). In order to prevent the polymerase from binding to chromosomal and plasmid DNAs contaminating the lysate, 5 M NaCl was added dropwise to the heat treated, clarified lysates to bring the final NaCl concentration to 0.25 M. DNA was then precipitated from this mixture by dropwise addition of 5% polyethylimine (in 20 mM TRIS•Cl, pH 8.5) to make the final concentration of PEI 0.3%. Precipitation was allowed to continue for 5 minutes on ice. A white, cloudy precipitate was removed by centrifugation at 15,000 x g for 15 minutes at 4°C. The supernatant fluid was decanted and saved. Following centrifugation, the NaCl concentration was reduced to 0.13 M by monitoring conductivity of the solution during the addition of TETT minus NaCl (20 mM TRIS•Cl, 0.1 mM EDTA, 0.05% Tween-20, 0.05% Triton-X100, 1% glycerol, pH 8.5).

Excess PEI was removed using a Bio-Rex 70 (BIO-RAD, Richmond, CA) column (2.5 x 30 cm). The column was poured and equilibrated with TETT Buffer + 0.1 M NaCl. The polymerase does not bind to the Bio-Rex 70 under these conditions.

To remove contaminating *E. coli* proteins, the Bio-Rex 70 column eluate was loaded directly onto a Heparin-Agarose (Sigma Chemical Company, St. Louis, MO) column (1.5 x 30 cm) which was also poured and equilibrated in TETT Buffer + 0.1 M NaCl. The heparin-agarose column was washed with 2 column volumes of TETT + 0.1 M NaCl and Taq DNA polymerase was eluted as a sharp peak using TETT + 1 M NaCl. Elution was monitored at 280 nm.

The heparin-agarose column fractions corresponding to the peak absorbance were pooled and concentrated to 0.15 ml using Ultrafree-15 Centrifugal Filter Devices (Millipore Corporation, MA) according to the manufacture's recommendations for centrifugation speeds and times. The concentrate was diluted to 15 ml with TETT Buffer + 5% glycerol and the sample was re-concentrated to 0.15 ml. This was repeated one more time to reduce the final NaCl concentration below 1 mM in the protein samples.

The concentrated polymerase samples were diluted two-fold using TETT + 5% glycerol and an equal volume of TETT + 95% glycerol was added to bring the final glycerol concentration to about 50%. Samples were stored at -20 °C. Protein concentrations were determined using the "Bradford Protein Assay" (BIO-RAD, Richmond, CA). Activity was measured using a radiometric assay (described elsewhere).

Typical yields of polymerase from 2-liters of induced *E. coli* culture (corresponding

to 30- 50 ml of heat treated, clarified lysate) ranged from 4 to 24 mg. SDS-PAGE analysis of the purified samples showed one dark band of about 94,000 molecular weight) and several minor ones after Coomassie Blue staining. The gels indicated a typical purification level of > 90%.

Example 2

Selectivity Assay

An unlabeled versus dye-labeled terminator assay ("terminator" is defined as a non-extendible base such as 2',3'- ddNTPs) was used to screen mutant Taq DNA polymerase samples for better Tet(II)•ddCTP incorporating mutant forms of this polymerase. This assay is based upon two substrates competing for the same active site at the same time during a steady state reaction in which only the polymerase concentration is limiting. Therefore, the assay measure the polymerase's "selectivity" for the unlabeled versus the fluorescein-labeled terminator. The DNA Primer/template used in this assay format is given below:

5'->(FAM)-CCC TCG CAG CCG TCC AAC CAA CTC A

GGG AGC GTC GGC AGG TTG GTT GAG T**G**C CTC TTG TTT<-5'

The next template position following the 3'-end of the primer is indicated above by the bold and underlined G.

The reaction consisted of:

80 mM TRIS•Cl (pH 9.0 at 20 °C)

1000 nM DNA primer/template [5'-(FAM)25mer / 36 G_1 template]

2 mM $MgCl_2$

5 0 µM TET(II)•ddCTP

1 µM ddCTP

0.25 Units of enzyme

40 µL reaction volume

60 °C reaction temperature

Samples (2 µL) were removed from the reaction mixture at pre-determined times (typically, 20 second intervals for 0.25 Units of polymerase activity per µL) and added to ice

cold 50 µL 0.5 M EDTA (pH 8.0). Timed aliquots were mixed and held on ice for further processing.

Samples of each time point were processed to remove excess, unincorporated TET(II)•ddCTP. Typically, 1.6 µL of each quenched sample were added to 250 µL of 0.8 M LiCl plus 0.2 µg/ml *E. coli* tRNA. followed by 750 µL of 95% ethanol. After mixing, the nucleic acids were allowed to precipitate for 20 minutes at -20 °C. The precipitates were recovered by centrifugation using standard procedures. The supernatant fluid was discarded and pellets were dissolved in 50 µL of 50% formamide. Gel samples were heat treated (95°C for 2 minutes) and 2 µL were loaded per sample lane on a 16% denaturing DNA sequencing gel. Gels were run on an Applied Biosystems Model 373 Sequencer using GeneScan Fragment Analysis software to measure the amount of FAM fluorescence in the bands corresponding to the 25-mer primer, the 26-mer product (indicating a ddC incorporation event) and the apparent " 27-mer" product band (indicating a TET(II)•ddC incorporation event).

The fluorescence signal in each of the bands was summed and the percent of signal in each band was used for further calculations as a normalization to avoid lane to lane loading differences. Energy transfer from the %-FAM moiety present on the apparent "27-mer" product molecules to the Tet(II) moiety on the newly incorporated 3'-base was not corrected since all ratios were compared to "wild type" or Taq G46D.) The normalized fluorescent signals in the 26-mer and "apparent" 27-mer product bands were corrected for the different concentrations of the two molecules used in the reaction and the corrected values were plotted versus time. The velocity of incorporation for each substrate was determined using least square fits to the data. The ratio of ddC / TET(II)•ddC incorporation rates is equal to the selectivity bias that the sample polymerase shows for the unlabeled versus the TET(II)-labeled nucleotides and reflects the following relationship:

$$v_{ddC} \div v_{Tet\bullet ddC} = (k_{cat} / K_M)_{ddC} [\, ddC \,] \div (k_{cat} / K_M)_{Tet\bullet ddC} [\, Tet(II)\bullet ddC \,]$$

where:

v_{ddC} = velocity of ddC incorporation

$v_{Tet(II)\bullet ddC}$ = velocity of Tet(II)•ddC incorporation

k_{cat} = catalytic rate constant

K_M	=	nucleotide equilibrium binding constant
[ddC]	=	concentration of ddCTP in the reaction
[Tet(II)•ddC]	=	concentration of Tet(II)¥ddCTP in the reaction

In this assay format, "wild-type" Taq or (Taq G46D) showed a selectivity bias or ddC / Tet(II)•ddC number of about 85 to 1. Mutants showing lower selectivity bias ratios were submitted to further testing. The Table 2 below shows the results for a few of the mutants tested by way of a few examples:

Table 2

Taq	Selectivity Number	WT / Mutant
G46D	85	85 / 85 or 1
G46D; R660D	8	85 / 8 or ≈ 10
G46D; R595E	28	85 / 28 or ≈ 3
G46D; F667Y	28	85 / 28 or ≈ 3
G46D; E681G	40	85 / 40 or ≈ 2
G46D; D655L	40	85 / 40 or ≈ 2

Example 3

Next Nucleotide Rate Effect Assay

An additional kinetic step between "ground state" nucleotide binding or initial collision and correct base pair formation and the group transfer reaction would be expected to slow the polymerase dissociation rate from an Enz•DNA complex having a 3'-dideoxynucleotide in an assay termed the "Next Nucleotide Rate Effect" (Patel et al., 1991). This assay measures the steady state rate of incorporation of ddTTP (i.e., the enzyme is limiting) in the absence or presence of the next correct nucleotide. The primer template pair is shown below:

5'->(FAM)-CCC TCG CAG CCG TCC AAC CAA CTC A

GGG AGC GTC GGC AGG TTG GTT GAG T**A** GTC TTG TTT<-5'

The next template position is indicated by the bold, underlined A. The next template

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

position beyond A is G. Under steady state reaction conditions, essentially all of the available polymerase is bound to the primer/template. When ddTTP is present alone in solution, it is incorporated following binding to its template position, A. Additional incorporation events require the polymerase to dissociate from the Enz•DNA complex and find another available primer/template that has not already undergone and incorporation event. Hence the rate of incorporation under these conditions is the dissociation rate of the polymerase from the Enz•DNA complex. If the next correct nucleotide, dGTP or ddCTP, is also present in the reaction mixture, the dissociation rate of the polymerase from the Enz•DNA•ddCTP complex, for example, will be slower if there is an additional kinetic step between the group transfer reaction that incorporated the ddTTP and an attempt by the polymerase to incorporate ddCTP in a processive mode of synthesis. This slower rate of dissociation an be detected as a slower incorporation rate of ddTTP since no chemistry can occur once ddTTP and the polymerase can no be processive despite the presence of another correct nucleotide. As shown in Figure 2, the presence of the next correct nucleotide does indeed slow the turnover or dissociation rate of the polymerase (Taq G46D; F667Y). Figure 2 also shows that the presence of a fluorescein dye on the next correct nucleotide (in this case, Tet(II)•ddCTP), appears to accelerate the turnover rate. We interpret this to mean that the polymerase is constantly undergoing a conformational change and that it can attempt to undergo the change even in the absence of the next correct nucleotide. However, the presence of a fluorescein dye on the next correct nucleotide blocks the ability of the polymerase to undergo such a change and thereby causes an immediate dissociation of the enzyme following the group transfer step for ddTTP incorporation. Hence, the fluorescein dye appears to accelerate the polymerase dissociation rate by eliminating a kinetic step (or steps) following the group transfer reaction.

Figure 3 shows the results for a Next Nucleotide Rate Effect assay for a "multiple" mutant form of Taq DNA polymerase, Taq G46D; R660D; F667Y; E681G. In this case, the presence of Tet(II) on the next correct nucleotide is "transparent" to the mutant polymerase. We interpret this to mean that the mutant polymerase can indeed undergo the same kinetic steps following group transfer that "wild-type" versions of this polymerase undergo. We also interpret these results to indicate that the F667Y mutation belongs in a different class than the R660D or E681G mutations since Taq G46D; F667Y still shows a "fluorescein-effect" in the "Next Nucleotide Rate Effect" assay, however, the multiple mutant, Taq G46D; R660D, F667Y; E681G, does not.

Typical assay conditions for the Next Nucleotide Effect assay were as follows:

1000 nM primer/template DNA
80 mM TRIS•Cl (pH 9.0 @ 20° C)
2.4 mM $MgCl_2$
0.02 Units/μL polymerase activity
400 μM each nucleotide (when present)

Samples were taken and processed in the same manner as described under "Selectivity Assay." In this case, it is possible to distinguish a ddC-incorporation event from a Tet(II)•ddC incorporation event by the migration rate of the resulting fragments in a 16% gel. Incorporation of ddC results in a "normal" 26-mer band that migrates as expected above or slower than the 25-mer primer. Incorporation of Tet(II)•ddC results in slower migration causing the band to migrate with an apparent size equivalent to a 27- or 28-mer.

Example 4

Analysis of Additional Mutants

Table 1, provided below, provides a summary of results obtained with selectivity assays performed with several different Taq mutants. The analogous site for the mutation in the enzymes E. *coli* DNA polymerase I and phage T7 DNA polymerase are also noted. The term "FS" refers to a Taq DNA polymerase having a F667Y mutation.

References

Barnes, W.M. (1992) The fidelity of Taq polymerase catalyzing PCR is improved by an N-terminal deletion. *Gene* 112: 29-35.

Brandis, J.W., Edwards, S.G. and Johnson, K.A. (1996) Slow rate of phosphodiester bond formations accounts for the strong bias that Taq DNA polymerase shows against 2',3'-dideoxynucleotide terminators. *Biochemistry* 35: 2189-2200.

Desai, U.J. and Pfaffle, P.K. Single-step purification of a thermostable DNA polymerase expressed in Escherichia coli. *Biotechniques* 19: 780-784.

Fersht, A. (1985) in "Enzyme Structure and Function," W.H. Freeman and Company, 2nd ed., pp. 111-112.

Johnson, K.A. (1993) Conformational coupling in DNA polymerase fidelity. *Ann. Rev. Biochem.* 62: 685-713.

Patel, S.S., Wong, I., and Johnson, K.A. (1991) Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. *Biochemistry* 30: 511-525.

Incorporation by Reference

This application incorporates all publications, patents, and patent application referenced herein in there entirety.

Equivalents

While the invention has been described and illustrated with reference to specific embodiments, those skilled in the art will recognize that modifications and variations may be made without departing from the principles of the invention as described hereinabove and set forth in the following claims.

No.	Taq Mutant	Pol I Equiv.	T7 Equiv.	[Final] µg/µl	*Units per µl	Spec. Act.	TET(II)ddCTP/ ddCTP (Mut / Wt)	ROX•ddCTP/ ddCTP (Mut / Wt)	TAMRA•dTTP/ ddTTP (Mut / Wt)
1	G46D			4.4	200	45.5	1 *	1	1
	G46D (LS.1)			28.0	1950	69.6			
	G46D (LS.2)			78.3	6240	79.7			
	FS			RMS			3		920
2	G46D; F667W	F762	Y526	3.6	5	1.4	1	1	
	(LS.1)	F762	Y526	25.0	150	6.0			
3	G46D; R573E	R668	R429	11.0	0	0.0			
4	G46D; E615L	E710	E480	4.8	0	0.0			
5	G46D; E615D	E710	E480	10.4	600	57.7	no activity		
6	G46D; E615I	E710	E480	7.0	140	20.0	no activity		
7	G46D; R587K	R682	"V443"	8.9	420	47.2	1	1	
8	G46D; R573K	R668	R429	9.3	0				
9	G46D; L657T	E752	T517	9.5	450	47.4	1	1	
10	G46D; R587K	R682	"V443"	nd	nd				
11	G46D; Q754S	Q849	Q615	13.0	0	0.0			
12	G46D; E615K	E710	E480	1.8	0	0.0			
13	G46D; R573Q	R668	R429	14.0	800	57.1	1	1	
14	G46D; D655L	T750	L515	11.3	400	35.4	2	0.7	
15	G46D; Q754K	Q849	Q615	nd	nd	nd			
16	G46D; R595K	R690	H460	nd	nd	nd			
17	G46D; K831M	H928	H704	8.3	300	36.1	1	1	
18	G46D; L682G	L777	"1540"	5.5	200	36.4	1	1	6
19	G46D; R659K	R754	D519	22.9	150	6.6	no activity		
20	G46D; A683E	N778	V541	12.8	900	70.3	1		
21	G46D; Q754K	Q849	Q615	5.0	0	0.0			
22	G46D; R593H	R688	E458	16.7	700	41.9	1		
23	G46D; R595E	R690	H460	23.5	50	2.1	3		
24	G46D; A683V	N778	V541	11.4	340	29.8	1		
25	G46D; Q592A	R687	A457	nd	nd	nd			
26	G46D; R660D	R755	D519	13.3	190	14.3	10		
27	G46D; T640G	R735	T507	13.0	225	17.3	1		
28	G46D; E681G	Q776	"1540"	7.5	170	22.7	2		
29	G46D; V654E	V749	E514	9.2	210	22.8			
30	G46D; Q613E	Q708	G478	15.2	71	4.7			
31	G46D; Q610A	Q705	D475	16.6	0	0.0			
32	G46D; E820K	E917	E693	11.6	475	40.9			
33	G46D; L817A	L914	L690	15.2	470	30.9			
34	G46D; I684G	I779	G542						
35	G46D; R660D; F667Y			19.1	179	9.4	10		
36	G46D; R595D; R660D; F667Y			10.5	0	0.0			
37	G46D; D655L; R660D; F667Y			18.8	228	12.1	10		
38	G46D; R660D; F667Y; E681G			13.1	404	30.8	12		
39	G46D; R595E; F667Y			9.5	0	0.0			
40	G46D; T7 Loop JK			12.7	0	0.0			
41	G46D; A582QS83			13.2	0	0.0			
42	G46D; P656S	S751		16.2	560	34.6			
		K731							
	40 lysates				22 act. mutants		17 tested 4 > "wt"		
							* WT = 85		

Table 1

SUBSTITUTE SHEET (rule 26)

01/22/2003, EAST Version: 1.03.0002

R660- Mutants

	Lysate #	Genotype	Specific Activity	Tet Selectivity TET(II)•ddCTP / ddCTP * (Mutant / WT)
Acidic-				
Aspartic acid	29	CS; R660D	14	10*
	38	FS; R660D	9	10
	39	R595E; FS; R660D	0	nd
	40	D655L; FS; R660D	12	10
	41	FS; R660D; E681G	31	12
	49	CS; R660D	41	nd
Glutamic Acid	51	FS; R660E; E681G	11	7
	72	FS; R660E	1	7
Basic-				
Lysine	50	FS; R660K	28	1**
Histidine	101	FS; R660H	13	1
Imino-				
Proline	66	FS; R660P	8	1
Aliphatic-				
Alanine	68	FS; R660A	4	4
Isoleucine	73	FS; R660I	5	0.9***
Valine	90	FS; R660V	10	1
	55	FS; R660V; E681G	1	1
Leucine	91	FS; G660L	8	0.6***
	52	FS; R660L; E681G	28	1
Glycine	47	FS; R660G; E681G	18	6
	78	FS; R660G	8	2
Polar Uncharged-				
Glutamine	53	CS; R660Q	47	1
	69	FS; R660Q	5	3
Serine	98	FS; R660S	16	7
Cysteine	93	FS; R660C	14	4
Asparagine	97	FS; R660N	13	3
Threonine	96	FS; R660T	26	3
Methionine				
Aromatic-				
Phenyalanine	92	FS; R660F	9	0.1***
Tyrosine	95	FS; R660Y	17	1

Table 2

E681- Mutants

	Lysate #	Genotype	Specific Activity	Tet Selectivity TET(II)•ddCTP / ddCTP * (Mutant / WT)
Acidic- Aspartic acid	71	FS; E681D	9	4**
Basic- Lysine	75	FS; E681K	52	6
Arginine				
Histidine	86	FS; E681H	37	7
Imino- Proline	74	FS; E681P	19	9
Aliphatic- Alanine	63	FS; E681A	13	6
Isoleucine	99	FS; E681I	37	27
Valine	76	FS; E681V	110	10
Leucine	87	FS; E681L	22	14
Glycine	48	FS; E681G	37	6
Polar Uncharged- Glutamine				
Serine	61	FS; E681S	12	5
Cysteine	88	FS; E681C	20	2
Asparagine	89	FS; E681N	40	4
Threonine	81	FS; E681T	35	6
Methionine	85	FS; E681M	32	47
Aromatic- Phenyalanine				
Tyrosine	80	FS; E681Y	42	3
Tryptophan	84	FS; E681W	37	17

11-05-97

*Ratio > 1 means improved TET(II)•ddCTP incorporation.
enzyme.
**Ratio = 1 means wild-type activity.
***Ratio < 1 means activity worse than wild-type.

M > I >W > L > V > P > H=K=G=T=S > D=A=N > Y=C
47 27 17 14 10 9 7 6 6 6 5 4 4 4 3 2

Table 2 (Continued)

01/22/2003, EAST Version: 1.03.0002

Tryptophan	94	FS; R660W	8	1

D > E=S > C=A=Q=T =N> G > K=P=V=Y=W=H > I = L >> F
10 7 7 4 4 3 3 3 2 1 1 1 1 1 0.9 0.6 0.1

*Ratio > 1 means improved TET(II)•ddCTP incorporation. Must be "85" to be "transparent to the enzyme.
**Ratio = 1 means wild-type activity.
***Ratio < 1 means activity worse than wild-type.

Table 2 (Continued)

CLAIMS

What is claimed is:

1. A DNA polymerase having at least one mutation in the nucleotide label interaction region, wherein the DNA polymerase has reduced discrimination for fluorescein-type dye labeled nucleotides.

2. A DNA polymerase according to claim 1 , wherein the mutation is in portion of the nucleotide-label interaction region selected from the group consisting of (i)the O-helix, (ii) the K helix, and (iii) the inter O-P helical loop.

3. A DNA polymerase according to claim 2, wherein the mutation is segment of the enzyme corresponding to amino acid residue selected from the group consisting of E520. A531, L522, R523, E524, A525, H526, P527, I 528, V529, E530, K531, I532, R536, E537, R573, Q582, N583, V586, R587, P589, Q592, R593, R595, D610, T612, Q613, E615, R636, D637, T640, F647, V654, D655, P656, L657, R659, R660, T664, E681, L682, A683, I684, P685, E688, F692, Q754, H784, L817, E820, L828, K831, and E832.

4. A DNA polymerase according to claim 3, wherein the mutation is at a position selected from the group consisting of, R595, D655, R660, T664 and E681.

5. A DNA polymerase according to claim 4, wherein the DNA polymerase is Taq DNA polymerase.

6. A DNA polymerase according to claim 5, wherein the mutation is selected from the group consisting of R660D, D655L, E618G, and R595E.

7. A DNA polymerase according to claim 6, wherein comprising a mutation set belonging to the group consisting of (G46D, R660D, F667Y), (G46D, R595D, R660D, F667Y), and (G46D, R660D, F667Y, E681G), and (G46D, F667Y, E681G).

8. A DNA polymerase according to claim 2, wherein the DNA polymerase is a thermostable DNA polymerase.

9. A polynucleotide encoding a DNA polymerase according to claim 1.

10. An expression vector having a promoter, wherein the vector comprises a polynucleotide according to claim 1 in functional combination with the promoter.

11. A host cell comprising an expression vector according to claim 10.

12. A method of synthesizing a fluorescently labeled polynucleotide, said method comprising the step of mixing a DNA polymerase according to claim 1 with a primed template.

13. A method according to claim 12, wherein the primed template is a primed template in a chain termination sequencing reaction.

14. A method according to claim 12, wherein the primed template is a primed template in a polymerase chain reaction.

15. A kit for fluorescently labeling a labeling a polynucleotide, the kit comprising a DNA polymerase according to claim 1 and a fluorescently labeled nucleotide.

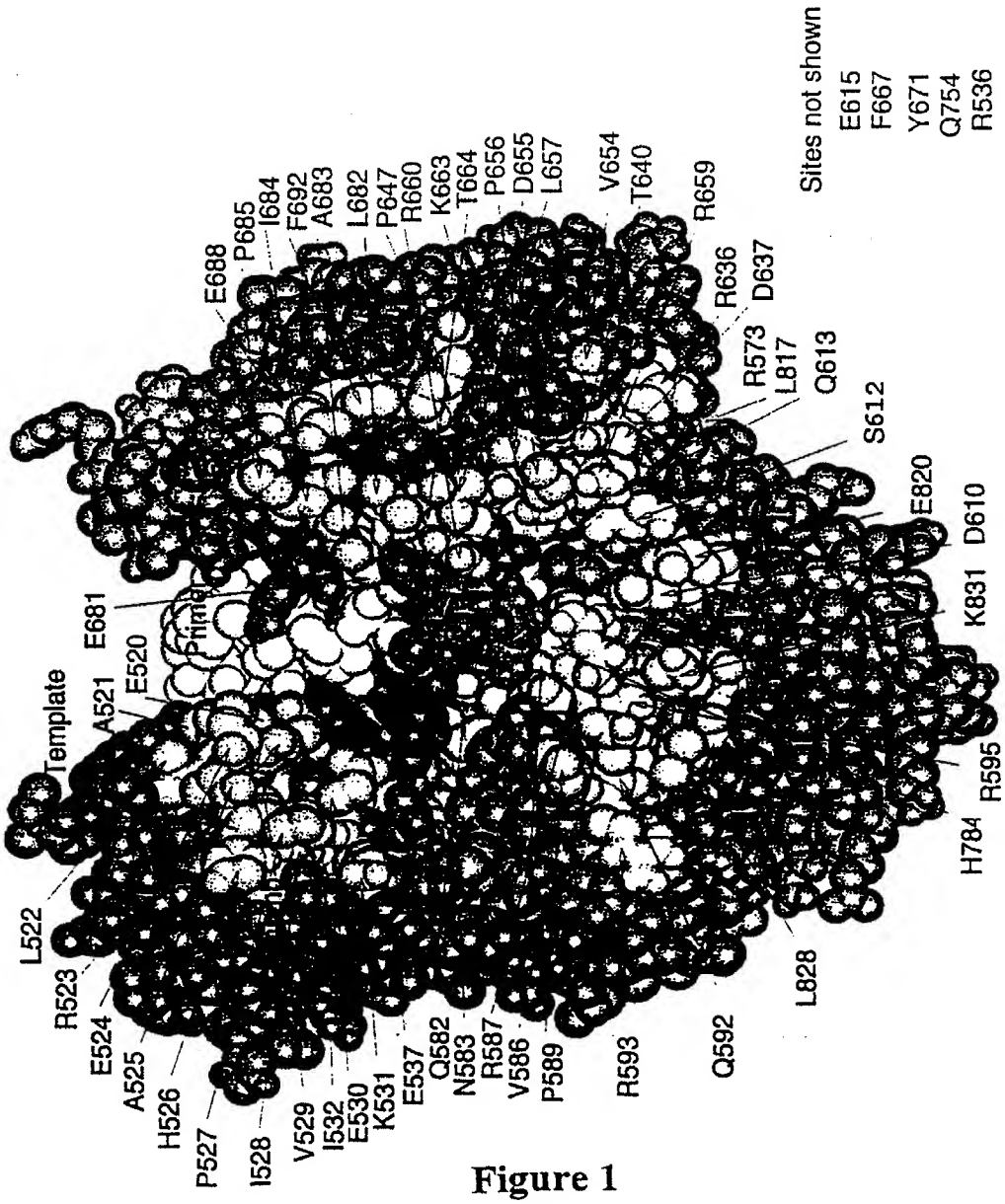


Figure 1

1/4

01/22/2003, EAST Version: 1.03.0002

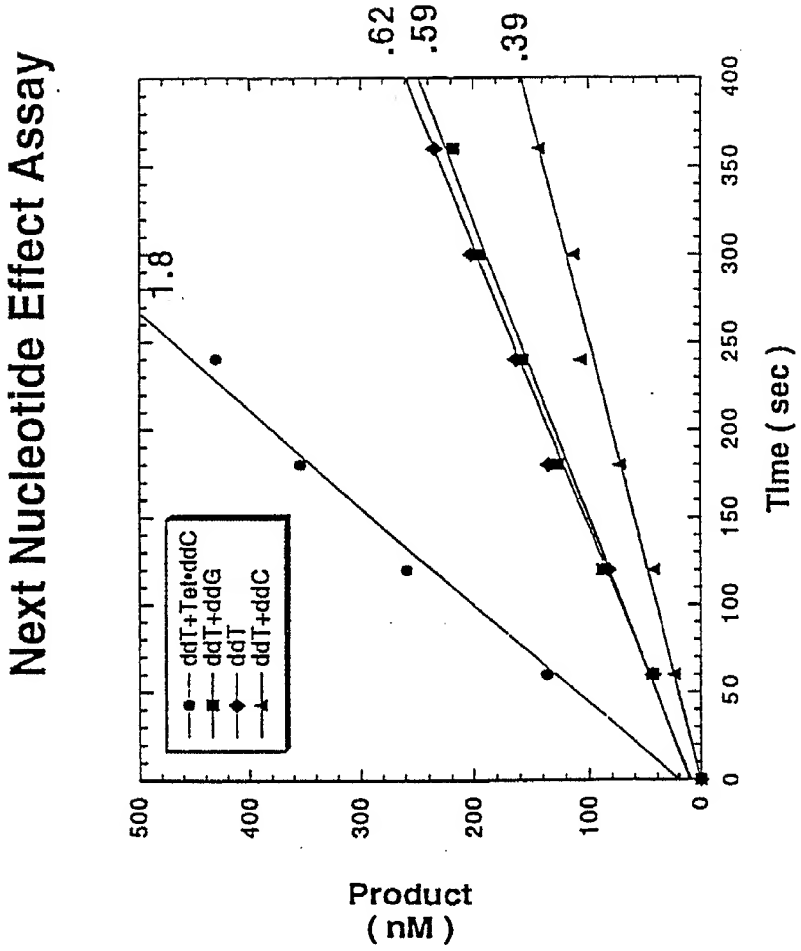


Figure 2

2/4

01/22/2003, EAST Version: 1.03.0002

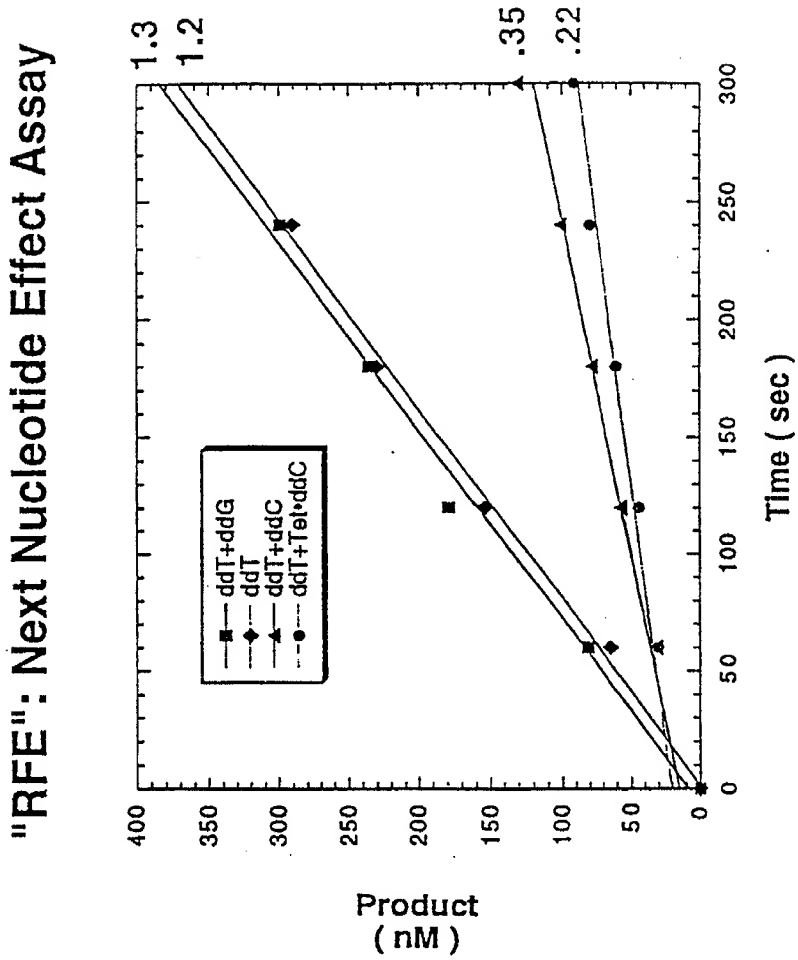


Figure 3

3/4

01/22/2003, EAST Version: 1.03.0002

TET-2 ddCTP

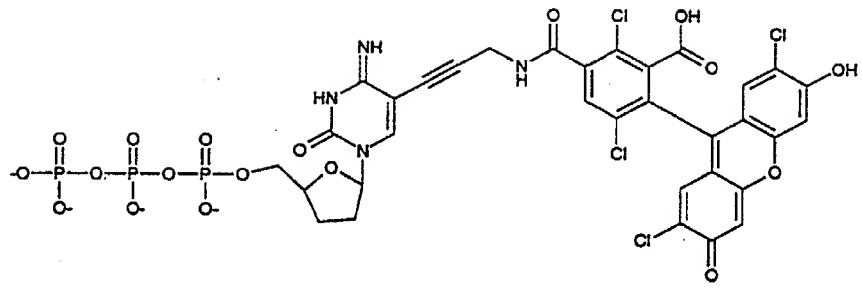


Figure 4

4/4

01/22/2003, EAST Version: 1.03.0002

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/05095

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/54 C12N9/12 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PARKER L. T. ET AL.: "AmpliTaq DNA polymerase, FS dye-terminator sequencing: analysis of peak height patterns" BIOTECHNIQUES, vol. 21, no. 4, October 1996, pages 694-699, XP002067706 see abstract; tables 1,2 see page 695, paragraph 2 - page 696, paragraph 1 see page 699, paragraph 2 -/--	1,2,8-15

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 June 1998	25/06/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Oderwald, H

Form PCT/ISA/210 (second sheet) (July 1992)

01/22/2003, EAST Version: 1.03.0002

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/05095

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	SUZUKI M ET AL.: "Random mutagenesis of Thermus aquaticus DNA polymerase I: concordance of immutable sites in vivo with the crystal structure" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE USA, vol. 93, September 1996, pages 9670-9675, XP002067707 see abstract; figures 2-4 see page 9671, paragraph 4 - page 9672, paragraph 6 ---	1-5,8-11
X	REEVE M A ET AL: "A NOVEL THERMOSTABLE POLYMERASE FOR DNA SEQUENCING" NATURE, vol. 376, 31 August 1995, page 796/797 XP000606193 see the whole document ---	1,2,8-15
X	EP 0 727 496 A (HARVARD COLLEGE) 21 August 1996 see abstract; figures 3,6; examples 1,7,11,20; table 2 ---	1,2,8-15
X	KALMAN L. V. ET AL.: "Thermostable DNA polymerases with altered discrimination properties" GENOME SCI TECHNOL, vol. 1, 1995, page 42 XP002067708 see abstract A-14 see the whole document ---	1,2,8-15
P,X	EP 0 823 479 A (HOFFMANN LA ROCHE) 11 February 1998 see abstract; examples 1,6; table 1 see the claims see page 8, paragraph 3 see page 2, line 30 - page 4, line 4 ---	1-3,8-15
P,X	VOSS H ET AL: "AUTOMATED CYCLE SEQUENCING WITH TAQUENASE TM: PROTOCOLS FOR INTERNAL LABELING, DYE PRIMER AND "DOUBLEX" SIMULTANEOUS SEQUENCING" BIOTECHNIQUES, vol. 23, no. 2, August 1997, pages 312-318, XP000698417 see the whole document -----	1,2,8-15

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

01/22/2003, EAST Version: 1.03.0002

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 98/05095

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0727496 A	21-08-1996	US 5614365 A	25-03-1997
		DE 29513622 U	19-10-1995
		DE 29513639 U	19-10-1995
		JP 2691149 B	17-12-1997
		JP 8205874 A	13-08-1996
		AT 143057 T	15-10-1996
		AU 4193396 A	06-05-1996
		CA 2201885 A	25-04-1996
		DE 69400567 D	24-10-1996
		DE 69400567 T	06-02-1997
		DE 655506 T	28-09-1995
		DK 655506 T	10-03-1997
		EP 0655506 A	31-05-1995
		ES 2072238 T	16-07-1995
		FI 971611 A	13-06-1997
		HU 77037 A	02-03-1998
		WO 9612042 A	25-04-1996
		ZA 9508761 A	13-05-1996
EP 0823479 A	11-02-1998	AU 3319797 A	12-02-1998
		JP 10066588 A	10-03-1998
		NO 973595 A	09-02-1998
		PL 321478 A	16-02-1998

01/22/2003, EAST Version: 1.03.0002